The Dissertation Committee for Sahar Nadeem certifies that this is the approved version of the following dissertation:

INFORMATION ACQUISITION IN NAVIGATION

Committee:

Brian J. Stankiwicz (Supervisor)

Mary Hayhoe (Co-Supervisor)

Lawrence Cormack

Benjamin Kuipers

Todd Maddox

INFORMATION ACQUISITION IN NAVIGATION

by

Sahar Nadeem, B.S.; M.A.

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

August 2008

ACKNOWLEDGMENTS

I would like to thank Dr. Brian J. Stankiewicz and Dr. Mary Hayhoe for their guidance and support during the duration of this project and for the careful and thoughtful reading of the drafts. This project was supported by an AFOSR grant FA 9550 awarded to Dr. Brian J. Stankiewicz.

INFORMATION ACQUISITION IN NAVIGATION

Sahar Nadeem, PhD.
The University of Texas at Austin, 2008

Supervisors: Brian J. Stankiewicz and Mary Hayhoe

The retention and recognition of landmarks within large-scale spaces (buildings or cities) plays an important role in way-finding and localization abilities. The current studies investigate our capacity for storing these views and the strategies used in deciding what information is stored and used. To investigate the issue of capacity we trained and tested subjects in six different environments with different levels of complexity. This manipulation was achieved by varying the number of states (position and orientations) within the environment from 10 to 132 in which each state generated a unique view. This manipulation generated environments in which the information content varied from 3 bits to 7.04 bits. We found no evidence of a capacity limitation for up to 7 bits of information. However, we did find that humans consistently lose about 1.25 bits of information regardless of the size of the environment. This finding was consistent in both virtual realty and in real environment. We further studied the nature of the information loss. Can gaze patterns reveal what information is being lost during the encoding process?

Table of Contents

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1. GENERAL INTRODUCTION

Humans possess the remarkable ability to learn highly complex environment's allowing them to navigate with ease and efficiency. They are able to complete both familiar and novel routes relying entirely on an internal memory representation of the space. There are several different approaches for representing space. For example, Siegel and White (1975) referred to the mental image of a large scale space as a *cognitive map* where information is acquired through a sequence of landmark, survey and route knowledge. Tversky (1993) preferred the term c*ognitive collage* and argued that the internal representation was not a single, cohesive map, but instead it was composed of multiple mini maps from varying sources of information that were woven together in a collage-like representation. Still others have argued that there is no need for an internal spatial representation. Instead they argue for a *view-graph* representation in which a view is associated with a specific action and goal (Schoelkopf & Mallot, 1995; Franz, Schoelkopf, Mallot, & Buelthoff, 1998).

Regardless of the type of representation proposed, an autonomous navigation system must make use of internal (e.g. proprioceptive) and/or external (e.g. landmark identification) cues for localization and navigation through large-scale spaces. Current research is interested in understanding whether the internal representation uses all of the information available in the real world or is there a limitation in the amount of information transferred?

The current research investigates the knowledge transfer limitations for large-scale spaces. That is, how much information from a large-scale space is being transferred to the mental representation of the environment formed by the individual? When traveling in a large-scale space, like walking to the grocery store from your apartment, all the cues that might aid navigation are not viewable from one vantage point. This leads to the need for the development of an internal representation of the environment that can be used to come up with direct routes towards goal positions that may not be visible from the starting point. The mental representation of the environment should also have the ability to recognize landmarks or routes from different perspectives making it flexible to use.

The real world is full of information and the conducted studies look at whether there are any capacity limitations for the transfer of information to the mental representation of space. If there are capacity limitations, what are the strategies that the human navigation system uses for selecting the information to be stored in memory? Does the position of unique landmarks in a large-scale space affect how they are used? Learning the pattern of information acquisition can help city or building planners decide where to place important information in large scale spaces to help people navigate. We can also learn how much information is required to impart accurate knowledge of a large scale space when giving directions.

1.1 INTRODUCTION TO NAVIGATION

Navigation skills play an important role in the survival of all mobile living animals. Animals use their navigational skills for vital tasks such as returning to a food source, returning to their shelters and migrating to warmer climates during winter and returning to food rich regions during the summer. Without the skill of navigation human beings, insects and animals would be unable to move from one known location to a specific goal without getting lost--making important acts like finding food, shelter and mates close to impossible.

Successful spatial navigation involves *localization* and *wayfinding*. *Wayfinding navigation* requires a general knowledge of the configuration of the environment in order to travel from one place to the next on a unique route. For example a person may know how to get home from work and how to get to the grocery store from home but has never had to navigate from work to the grocery store. Repeated path-following between home, work and store allows for the learning of the components of each route that can be recalled for later use. The integration of the individual routes into an internal representation can help build up short cuts and direct routes between places that you may not have earlier experience of navigating between (Golledge, 1999).

Localization is the process of specifying where in space (position and orientation) the observer is. Localization can be accomplished by relying on internal cues such as proprioceptive cues or external cues such as landmarks. *Path integration* also known as *dead reckoning* is a simpler form of localization that does not rely on external cues for localizing oneself in an environment. *Dead reckoning* is the ability to determine the change in one's position by calculating ones velocity with respect to time. Early sea farers depended on inaccurate astronomical cues and dead reckoning to guide their ships (Gallistel, 1990). Insects, birds and bees have been noted to be able to use *dead reckoning*. Desert ants wander meters away from their nests in a circuitous route to find food. However once the food is found the ant takes a direct route back to its nest (Gallistel, 1990). *Dead reckoning* relies on vestibular and proprioceptive cues rather than external landmarks like trees or rivers (Redish, 2000).

There are several different navigation strategies available that depend on external cues to localize one self successfully. *View-graph* is one navigation strategy that requires the use of external landmarks and cues. An individual navigates towards the goal state by comparing the stored view against the currently visible view. The goal is to match the view (associated with an action) saved in long term memory with the current view until the goal state is reached (Mallot, Franz, Scholkopf & Bulthoff, 1997). *Piloting* is another strategy that consists of walking away or towards a specific landmark. *Piloting* does require a mental representation of the environment to act as a reference frame for navigation. Hover flies are an example of insects that have been seen to use *piloting* to reach their home base (Gallistel, 1990).

In *route based navigation* each action is linked to a specific stimulus. It requires a memory for external cues like landmarks, turns, etc. Directions are usually given using *route based navigation*, for example, turn right at the second signal and then left at the gas station. *Route based navigation* does not require us to have a complete representation of the external environment as the person is only interested in getting to the 'second signal' and from there to the 'gas station'. You do not need to know where the 'signal' is in relation to the 'gas station' and other areas in the environment.

1.2 Cognitive Map

The term cognitive map was coined by Tolman (1948). He argued that the rat's ability to navigate in a maze was more then just a series of stimulus-response associations. Tolman, Ritchie and Kalish (1946), trained rats to run through an indirect path to a goal position that contained food. After the rats had learned the particular path, extra arms were added to the maze and the original route to the goal position was blocked. With this change in the maze rats were able to still choose either the most direct path to the food or pick a path running perpendicular to the food side of the room. The rats appeared to have learned more than just the initial path to the goal state. This view of a flexible internal representation was contrary to the behaviorist view of the time which saw the skill of navigation as a result of simply learning a sequence of responses to specific stimuli (Hull, 1943).

Siegel and White (1975) believed that the cognitive map of individuals developed with experience. Initially, in a new environment, people acquire elements from the real world like landmarks. Landmarks are salient and easily perceivable features present at specific locations in the environment (Stankiewicz & Kalia, 2007). Landmarks can act as the strategic focal positions at the beginning and ending of a path and they can also be used as intermediate course maintaining loci. However, with repeated traveling between the focal points a shift occurs from landmarks to learning the routes between landmarks. The individual routes between different loci points eventually integrate together into a more complete survey representation of space (Siegel & White, 1975).

Tversky (1993) proposed that people have several different sources of information, for example written cues, auditory instructions or active exploration and it is unlikely that the pieces of information can be organized into one, map-like mental image. People also show hierarchical representations by placing well known areas closer together and overestimating distances between lesser known targets. For example in a study by Stevens and Coupe (1978) students at U.C. San

Diego were asked to map Reno in relation to San Diego. The subjects tended to incorrectly state that San Diego was to the west of Reno which is inaccurate. Other factors like the individual's perspective depending on personal knowledge of a region or regularization of geographical features (people tend to think of the environment along symmetrical lines, regularizing the environment) further weaken the idea of a singular map-like cognitive map based on metric information.

Tversky (1993) supported the idea of a *cognitive collage*. A collage would contain information from all kinds of sources forming several mini maps that would come together. The internal representation would not be a coherent map-like structure as it would lack accurate metric information and thus it would have distortions due to many factors such as incomplete information or different perspectives.

In simpler environments, where people seem to have an accurate mental representation, the *spatial mental model* has been proposed (Tversky, 1993). This model integrates landmarks in a perspective free manner. The landmarks are isolated and learned in relation to one another, allowing for successful way-finding from different starting points and spatial inference. *Spatial mental models* would not have all the metric information but it would still represent coarse spatial relations among landmarks in a perspective free manner unlike the *cognitive collage* approach. Learning snapshots of images does not lead to successful navigation-- a change in view or perspective would not allow us to recognize the image. It is more efficient to learn landmarks in relation to one another allowing for successful navigation from various starting positions (Tversky, 1993).

The *view-graph approach* (Scholkopf & Mallot, 1995) is another theory for the acquisition of cognitive maps. According to this approach, humans store a series of views in memory. Each view is linked to a specific action. To reach the goal state a person sees a view and then recalls and executes the action associated with that view. After the first action is complete the next view comes into attention and the associated action is executed. This sequence would continue until the goal state is reached. According to Hull (1943), reinforcement helps learning and the immediate beneficial consequence of a particular behavior would increase the probability of the same behavior in the future, breaking down navigation into a series of learned responses to stimuli. The *view-graph approach* does not generate a complete survey representation of the environment and thus it is quite inflexible because each view has a predetermined action attached to it. The *view-graph approach* just focuses on each individual view and not the entire environment hence lacking a more holistic approach.

As the debate about how the internal spatial representation is learned and what information is explicitly represented in the map continues, most researchers do agree on the basic idea that there is some kind of mental representation that is used for successful navigation. In this paper any internal representation of an external space will be referred to as a "*Cognitive Spatial Representation*".

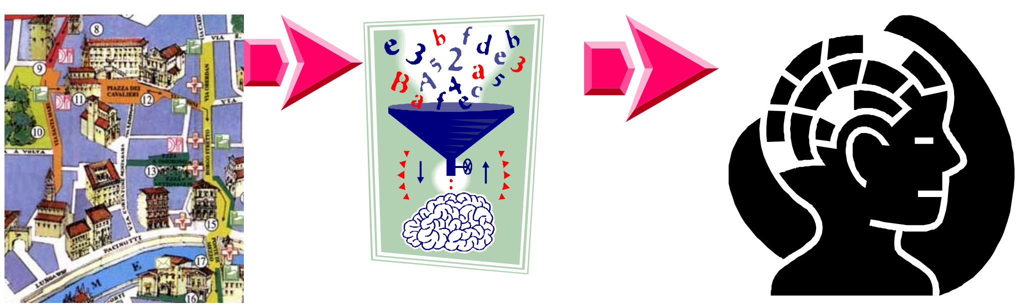
The various approaches to the formation of a *Cognitive Spatial Representation* have looked extensively at how the representation is generated and what is made explicit within the map. However, research has not looked at how much of the information available in a large scale space is being used and how that information is chosen. This paper looks at the nature of limitations for the amount of information that can be transferred from the real world to the *Cognitive Spatial Representation*. Current studies look at the general constraints of the *Cognitive Spatial Representation* and the strategies used to generate these representations.

CHAPTER 2. INFORMATION THEORY

Information theory provides a method for quantifying the amount of information in a stimulus and is primarily based upon the principles of uncertainty and variance. Furthermore, principles from *information theory* can be used to determine the capacity of a channel. It should be noted that anything that can transfer across space (e.g., telephone wire) or time (memory) can be considered a channel. Thus, the human cognitive system can be considered a channel and will be treated as such in the research conducted and proposed. More specifically, what are the capacity limitations of the human cognitive system with respect to the knowledge acquired about a large-scale space (memory) and how that information is transferred to the task of localization within that same large-scale space?

In the transfer of information there are three basic components: the source, the channel and the receiver. The cycle starts at the source which specifies the original information or message. The channel transfers the message from the source to the receiver. The receiver is where the information ends (Shannon, 1948). For example, when using a telephone the source is the person speaking into the telephone. The cables of the telephone and other miscellaneous pieces of equipment serve as the channel that carries the signal from the speaker to the other person. In this case, the person on the other end is the receiver. The three basic components exist in all fields where there is any flow of information e.g., radio, television, semantics, psychology etc.. Abundant research has been carried out in diverse fields such as mathematics (inequalities), physics (AEP thermodynamics), economics (Portfolio theory) that look at the flow of information according to the concepts of *information theory* (Cover & Thomas, 2006).

In the case of spatial navigation, the source would be the information available in the environment, and the channel would be an individual's senses and cognitive skills that interpret and transfer the available information. The *Cognitive Spatial Representation* formed by an individual for the specified environment would be the receiver as seen in Figure 1.



Source Channel Cognitive Spatial Representation

Figure 1: Information transfer in spatial navigation

For some channels, there is a loss of information when going from the source to the receiver. There could be several different causes for the loss of information. For example, a noisy channel is unable to carry all the information correctly. A noisy channel corrupts the information in transit. On the other hand if a channel is capacity limited there will also be information loss. The information would be transferred perfectly up to a certain limit after which no further information would be transferred. To measure the amount of information being transferred we can use principles of *information theory*.

In its simplest form, *information theory* measures the amount of uncertainty or the variance of the input pattern that goes in the channel compared to the variance that comes out of the channel and the correlation between the input and output patterns (Shannon & Weaver, 1949; Shannon, 1993; Miller, 1956). As the correlation between the source pattern and the received pattern increases, the channel is increasing the information transferred. However if the correlation between the source pattern and the received pattern decreases, then the channel is limited.

In information transfer the challenge is to transfer large amounts of data with the least amount of noise corrupting the information. For example, satellite signals to a television set are transferring large quantities of data at all times. There is a certain amount of noise that disrupts the data transfer but if the interruptions in the signal increase the image on the television screen looses its clarity and sharpness becoming grainy due to a larger loss in the amount of information being transferred. Parity bits or error correcting codes can be added to the original data to decrease the disruptions in the information transfer. However this can slow down the transfer of data from the source to the receiver. Information Theory as originally proposed by Shannon (1949) looks at how much data can be compressed and transferred without too much noise (Mackay, 2003). Information can be measured as uncertainty, variance or entropy. Entropy is the uncertainty of a single random variable and it is measured in bits for this paper (Cover & Thomas, 2006).

2.1 The Bit

In *information theory*, the amount of information is measured in bits. A *bit* is a binary digit and can only have two states, one or zero, on or off. A *bit* of information is the amount of information required to make a decision between two equally possible outcomes (Shannon, 1948, Miller, 1956; Reza 1994). If we have to decide whether a person is over the age of fifty or under and both alternatives are equally likely, one bit of information would be enough to specify the right answer. A *bit* has logarithmic base 2, so every time the equi-probable alternatives increase by a factor of 2, one more *bit* of information is required (Reza 1994, Miller, 1956).

2.2 Entropy

Entropy is the uncertainty of a single random variable (X). It is a measure of the information content (number of bits) required on the average to describe variable 'X'. 'H' represents entropy in equation 1.

$$H(X) = \sum_{x \in X} p(x) \log_2 \frac{1}{p(x)} \qquad (1)$$

Logarithm to base 2 is used for measuring entropy in this paper.

2.3 Conditional Entropy

Conditional entropy is the entropy of a random variable (X), given another variable (Y). That is the average uncertainty that remains about 'X' when 'Y' is known.

$$H(X|Y) = H(X,Y) - H(X) \qquad (2)$$

$$H(X,Y) = H(X) + H(Y|X) \qquad (3)$$

The reduction in uncertainty in variable 'X' due to another random variable 'Y' is called the mutual information (Mackay, 2003).

2.4 Mutual Information

The current studies are interested in evaluating the transfer of information between the external environment (source) and human behavior. One common method for measuring the capacity of this transfer (channel) is to measure *Mutual Information*. Mutual information (I) is a measure of the amount of information one random variable (X) contains about another (Y) as seen in Equation 2. I(X,Y) is the reduction in uncertainty of 'X' due to knowledge about 'Y' (Cover & Thomas, 2006). It shows the variability of response in stimulus 'X' correlated with the variance in 'X' given that 'Y' is known (Reza, 1994).

$$I(X,Y) = H(X) - H(X|Y) \quad [1] \qquad (3)$$

Mutual information can be seen as the information contained in one process 'X' minus the information contained in 'X' when process 'Y' is also known. H(X) (calculated as shown in equation 1) in mutual information is the uncertainty of a single process 'X'. H (X|Y) (calculated as shown in equation 2) measures the conditional entropy; it shows the entropy in 'X' conditional on 'Y'. It shows us how much uncertainty remains for 'X' when given information about 'Y'.

---

[1] $I(X,Y) = H(X) - H(X \mid Y) = \sum_{x,y} p(x,y) \log p(x,y)/p(x)p(y)$

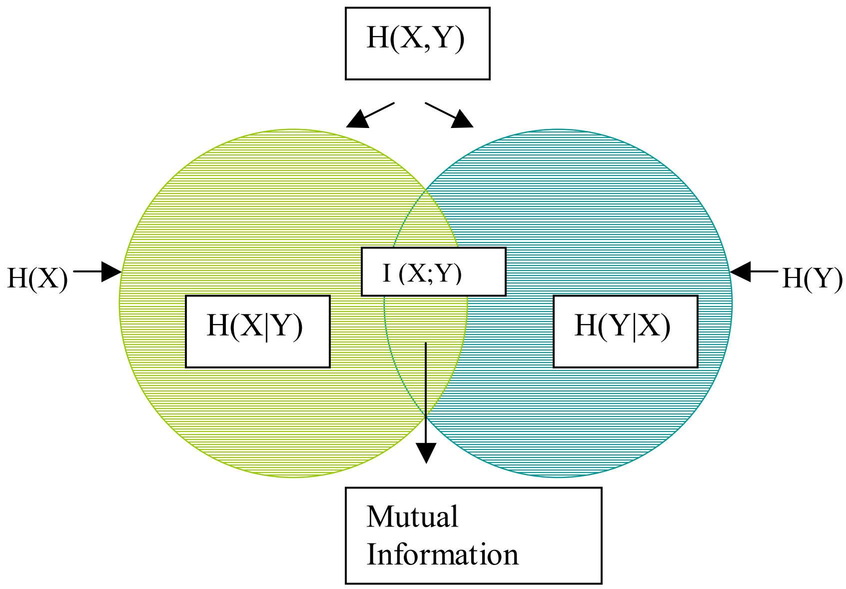
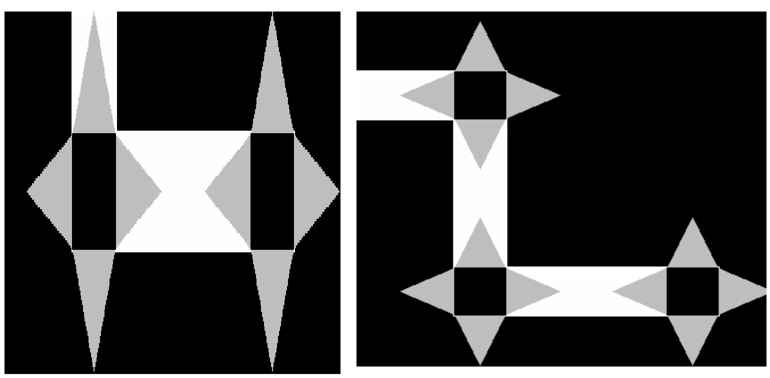


Figure 2: Venn diagram representation of Entropy, Mutual Information, Conditional and Joint Entropy (Reza, 1994).

2.5 Analyses in Research

Current work looks at the transfer of information from a spatial environment (source) through to the *Cognitive Spatial Representation* and to the behavior of the participant (specifically, in our case localization). Environments with varying amounts of information were used to look at the effect of increasing the information content on the transfer. The amount of information available in an environment was varied by increasing the number of states[2] (location and orientation) that had unique views within an environment as seen in Figure 3 (a,b,c). The smallest environment had only 2 corridors and the largest environment had 40 corridors. By increasing the size of the environments, the complexity of the environment was increased raising the bits of information available to a participant. If there is a capacity limitation then the increased number of bits available in an environment would not transfer to the *Cognitive Spatial Representation.* Mutual information allows for the bits of information being transferred to be measured.



[2] The triangles on the map as seen in Figure 3 are the specific positions and orientations in the environment that are tested.

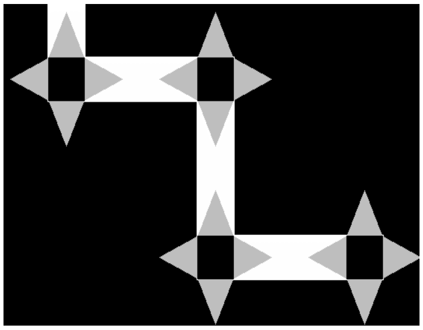


Figure 3a (upper left) a view of 1 corridor with 8 states and 3 bits of information.
Figure 3b (upper right) 2 corridors with 12 states and 3.5 bits of information.
Figure 3c (lower left) 3 corridors with 16 states and 4 bits of information.

If the channel is perfect, then the number of bits output would be the same as the bits of information available in the tested view as can be seen in the red line in Figure 4, however if there is a channel capacity then the channel would be perfect up to an extent and then level off at that specific number of bits as seen in the green line in Figure 4. Another possibility is of an imperfect channel in which only a percentage of the information would pass through no matter what the size of the input information is as seen in the blue line in Figure 4.

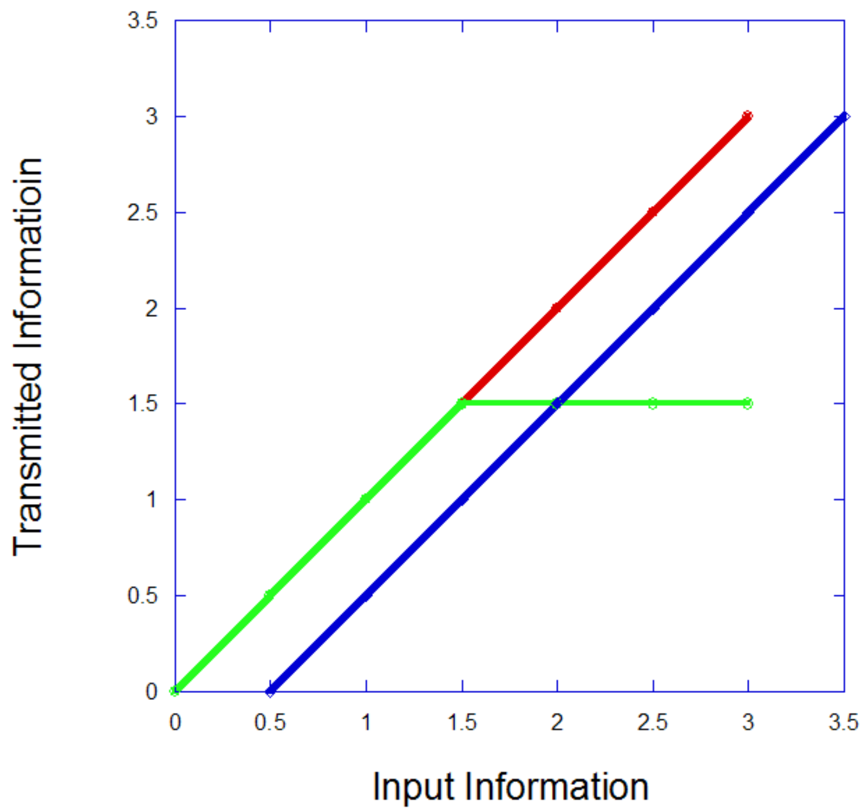


Figure 4: Three possible outcomes of how information is transferred from the real environment to the cognitive map.

To compute the capacity of the cognitive map the mutual information for each subject in each environment was calculated as a function of the given response (Y) and view (X) as shown in Equation 3. Mutual information tells us how well we can predict the response of a participant given that the view they are observing in the environment or what view the participant was shown given the participant's response. Through mutual information we can calculate the variance in responses and correlate it with the variance in responses given a specific view. It allows us to quantify the number of bits being transferred.

If every time a participant sees a particular view they always respond with the same position on the map, conditional entropy would be zero and the mutual information would equal to the overall entropy in response. On the other hand, if the given view does not remove any of the variance in the responses and the participant marks different positions in the repeated trials for the same view then the conditional entropy value would be high. $H(X|Y) = 0$, only if the value of 'X' is completely determined by the value of 'Y'. $H(X|Y) = H(X)$, only if the response (X) and view (Y) are independent random variables.

Mutual information results are sensitive to the consistency of responses but it does not tell us about the accuracy of responses. Therefore accuracy of responses was also calculated.

CHAPTER 3. EXPERIMENT ONE

In the field of cognitive psychology Miller's (1956) article, *The Magical Number Seven, Plus or Minus Two: Some Limits on our Capacity for Processing Information,* showed a pattern for the processing of uni-dimensional stimuli in working memory. Miller showed that when processing uni-dimensional stimuli there was a capacity limitation of around 2.8 bits (7 items)[3].

Experiment One investigated the limitations of information transfer from the real world to the *Cognitive Spatial Representation*. To study this we tested six participants in six virtual desktop environments that differed in size. The smallest environment had only two corridors and the largest one had forty corridors. "Pictures" were placed throughout the environment such that every state within an environment generated a unique view (as shown in Figure 6) and none of the pictures were repeated in the same environment and each picture was easily visible. Figure 5 shows examples of some of the environment maps, the triangles on the map show the states (position and orientation) that were tested in the experiment. The tested states were increased in each environment varying the bits of information available. In

[3] $2^{2.8}$=7items

the current study we varied the information from 3.58 bits ($2^{3.58}$=12 tested states) to 7.04 bits ($2^{7.04}$=132 tested states).

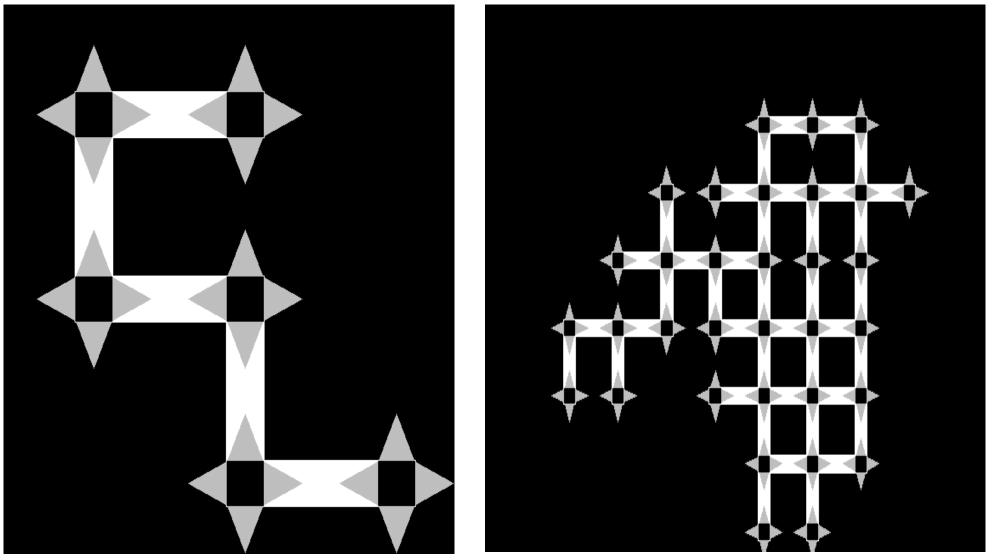


Figure 5: Map of a 5 corridor (left) and a 40 corridor (right) environment

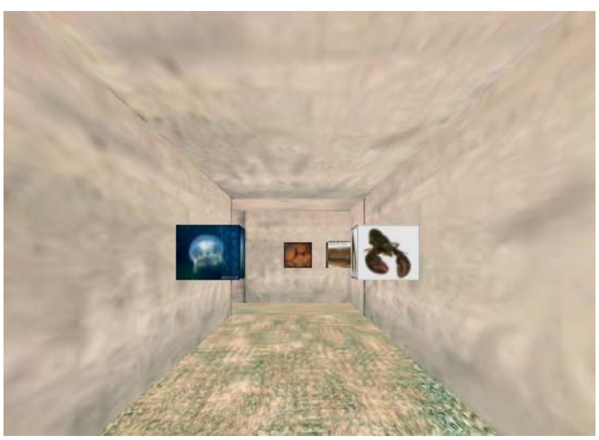


Figure 6: A view of a corridor with landmarks from the virtual environment

3.1 Procedure

Five undergraduate students from the University of Texas at Austin were paid $10 per hour to participate in the current experiment. We had 4 female participants' of an average age of eighteen. We had one male participant aged eighteen. The participants used Dell desktop computers for the experiment. Each participant moved through the environment using specific key presses on a normal keyboard. Pressing '8' on the keyboard moved the subject forward by one corridor unit. Key '4' rotated the participant 90 degrees counter-clockwise whereas key '6' rotated the participant 90 degrees clockwise. Each participant had to go through a training phase, testing phase and the experimental phase for each environment. Each subject ran through the six environments in a unique order. Table 1 shows the order in which the six participants ran in the study. Two different versions of each size environment were used (version A and B). The two versions had the same number of corridors but different topology (see Figure 7 for an example). The pictures used as landmarks in both versions were also different. All environments were run an equal number of times across subjects.

	Subjects				
Order	1	2	3	4	5
Of	5A	5B	2B	15A	25A
Environments	25A	2A	5B	10B	2A
	2A	15B	10A	25B	15B
	10B	10A	15B	2B	5B
	15A	25B	25A	5A	10A
	40B	40B	40A	40A	40A[4]

Table 1: Chart of random testing order

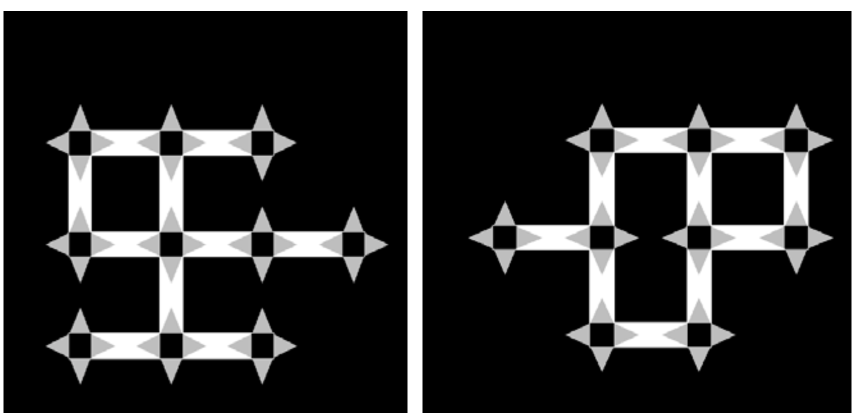


Figure 7: Map of environment 10A and 10B

---

[4] 40 corridor environment was added to the experiment later hence each participant ran in the environment at the end.

Environments were generated on a Cartesian grid such that hallways intersected at 90 degree angles. The height of each corridor from the ceiling to the floor was 1 meter with a width of 1 meter and wall thickness of .05 meters. Length of each corridor was 3 meters. The texture of the ceiling and walls was the same (cement) whereas the floor had a different texture (burlap) as seen in Figure 6. These environments were rendered using Vizard software. Participants viewed the environments from a first person perspective as seen in Figure 6.

In Experiment One we also ran seven more subjects (4 females and 3 males) who completed each environment with fog[5]. In the fog condition the participants had a limited view visibility was restricted to the length of one hallway unit. We carried out the fog condition to see if the transfer of information was affected by restricted visibility. Fog also allows us to determine the role that distal cues play during localization. The participants were students at the University of Texas at Austin and they were paid $10 per hour for their participation.

[5] In the fog condition participants did not complete the 40 corridor environment as it took too much time to learn such a large environment with fog.

Figure 8: View of a corridor with fog

Before starting the experiment the participants were told that the purpose of the study was to understand how information is acquired and used while navigating in a large scale environment. Participants were told that they were going to navigate through six different virtual environments of varying sizes.

3.1.1 Training Phase.

In the *training phase* participants started at a specified starting position in the environment and were told to explore and try to learn the entire environment and pay attention to its landmarks. The participants used the keys ‘8’, ‘4’ and ‘6’ to move through the environment. There were “hotspots” spread out all over the environments. The hotspots were positions in the maze that had a specific ID number and were tested during the *testing* and *training phases*. When the participants moved to a hotspot an auditory cue stated the hotspot ID. The computer would verbally announce that they had reached one of these locations ("*Position One*"). There were 3, 5, 6, 8, 9 and 11 hotspots respectively going from the smallest (2 corridor) to the largest (40 corridor) environment. The number of hotspots varied according to the size of the environments to make sure that the participants actually explored all the corridors in each environment.

Each participant explored the environment a set number of forward moves in the training phase. The number of forward movements was dependent upon the size of the environment (5 times the number of corridors in the environment). Once the training phase ended the participant went through the *testing phase* to check how well the participant had learned the environment. If the participants were unable to pass the *testing phase* they would return to the *training phase* for more exploration.

3.1.2 Testing Phase

This phase tested the participant's knowledge of the environment. The participants started off at a random hotspot location in the environment. An auditory cue announced the starting location and also instructed them to go from the current position to a specified hotspot, (e.g. "Position Eight, *Go To Position*

Four"). Once the participant reached the specified target location they were instructed to move to another hotspot location. If a participant passed over any other hotspot on the way to the current goal the computer provided an auditory cue stating the hotspot id (e.g. "*Position Five*"). Participants were instructed to take the most direct route from hotspot to hotspot, showing their knowledge of the environment. The computer measured the efficiency for reaching the target state from the starting state by comparing the number of moves it took the subject to reach the goal state against the minimum number of translations required for the navigation. The subject was tested on 10 hotspot location and had to complete five consecutive tests in a row that were better than 80% efficiency[6]. If the participant did not pass then a message would appear asking the person to return to the *training phase* else they moved on to the *experimental phase.*

3.1.3 Experimental Phase

During the *experimental phase* the subjects were shown a specific view from the explored environment. When the participant was ready to respond they indicated this by pressing the space bar and were shown a map of the environment similar to that shown in Figure 9B. The participant's task was to mark the location

[6] We calculated the least number of moves required traveling from the specified hotspot to the next and then we compared it to the number of moves that it took the participant to travel from the specified hotspot to the next. We compared the two values and if the participant had less then 80% efficiency then we counted it as a failed attempt.

and orientation of where the view was located. They could click on any one of the little triangles on the map and it would turn red. The participants were only allowed to mark one spot at a time on the map after which they clicked on the 'Save Data' button and moved on to the next trial. Each state was tested ten times in a random order.

The participants were allowed to stop the trials in the *experimental phase* at anytime and continue the experiment at a later time. This was to allow the participants to take a break if they were getting tired as the larger environments could take a long time. When the subjects stopped the trials in the middle, they had to go through the *testing phase* again to make sure that they still knew the environment well enough before they resumed with the *experimental phase*. If they failed the test they had to go through the *training phase* again.

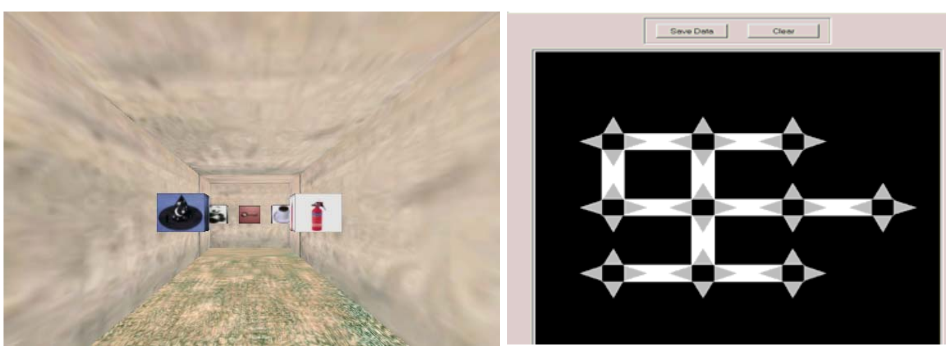


Figure 9A (left) is a view from the environment and Figure 9B on the right is a map of the environment

The participants also participated in a second *Experimental Phase* with fog which was also completed in the same way as stated above. The only difference was that the views that the participants observed had fog. The addition of the fog led to a limited view of the environment. The goal was to have an ability to determine the role of proximal versus distal landmarks along with isolating the role of particular landmarks within the environment.

3.2 Results

Mutual information was calculated to measure the number of bits being transferred from the environment to the Cognitive *Spatial Representation*. Mutual information was calculated for all participants for the different sized environments. Figure 10 shows the average mutual information across all participants. The green line shows the results from the participants in the fog condition who had a limited view of the environment and the red line shows the results from the participants in the no-fog condition.

We did not find a channel capacity for up to 7.04 bits or 132 states in the no-fog or the fog condition with a 100 states or 6.6 bits. However the channel was not perfect. There was a loss in memory of about 1 to 1.5 bits of information as the environments grew bigger. One bit may not seem like much of a loss, but it is equivalent to forgetting half of the states. Overall bit loss in each environment is shown in Figure 11.

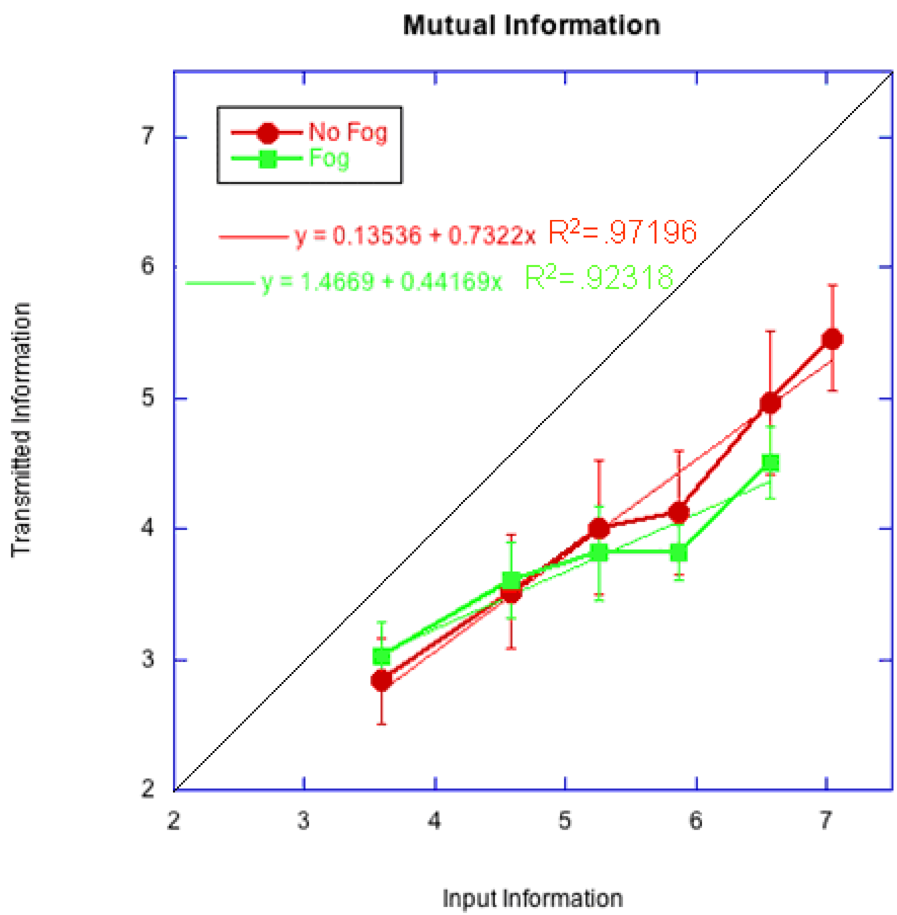


Figure 10: Experiment 1, Mutual Information: for each subject in the six environments. The green lines show the performance in environments with fog whereas the red line is the performance in environments without fog. Linear fit lines (red = no fog, green = fog) are also shown in this plot.

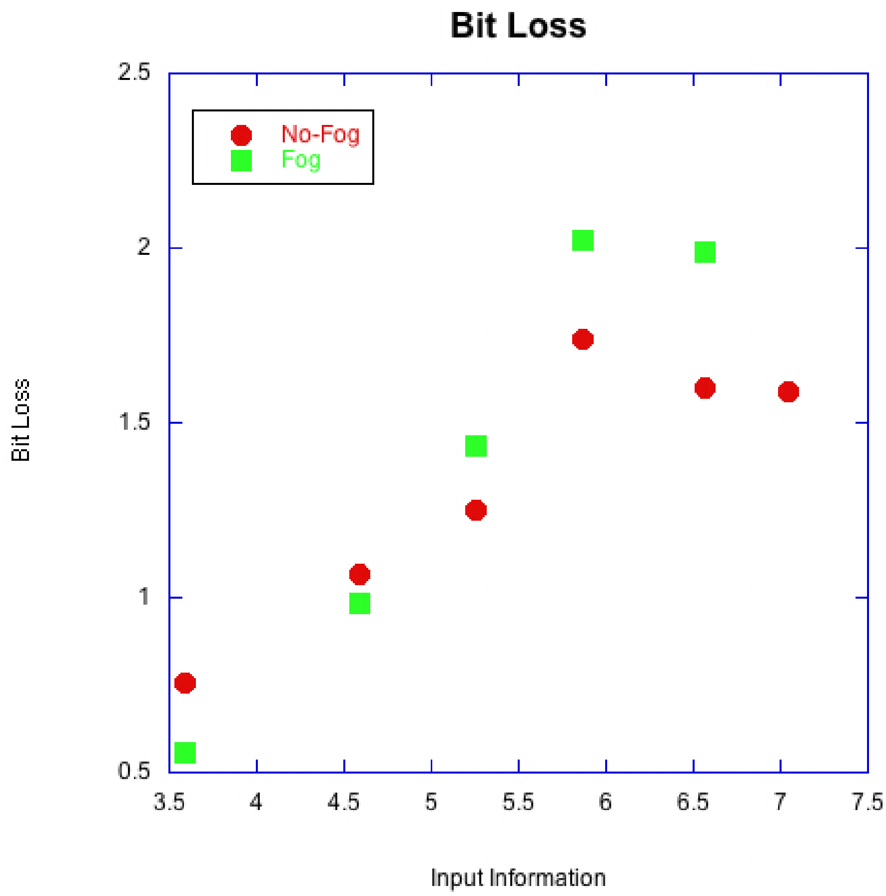


Figure 11: Experiment 1, overall bits lost in each environment. The green squares show bit loss in the fog condition whereas the red circles show the bits lost in the no-fog condition.

The fog environments help us see how the participants were using the cues in the environment. If they were using distal cues as aids, then, their performance should have declined in the fog condition as they could only see the immediate cues. However if the participants were primarily using the immediate cues the fog condition might actually be helpful as it removed all the distal cues from distracting the participants. In the results the limited view fog environments did not appear to have a significant detrimental affect on performance.

It appears that there is a limited capacity; people do not use all the information. We do not find a channel capacity; a linear fit is the best fitting line for the fog and no-fog results as seen in Figure 10. Information is not transferred up to a specific level and then asymptotes at that amount of information.

Mutual information results are sensitive to the consistency of responses. The ability to predict what view a participant was looking at given their response or vice-versa is given by the mutual information results. Accurate results will generate high mutual information values, but high mutual information values do not necessarily mean that the participant was accurate. Therefore, we also plotted the accuracy of responses for both the fog and no-fog condition as seen in Figure 12. The green line shows the results from the participants in the fog condition who had a limited view of the environment and the red line shows the results for the participants in the no-fog condition.

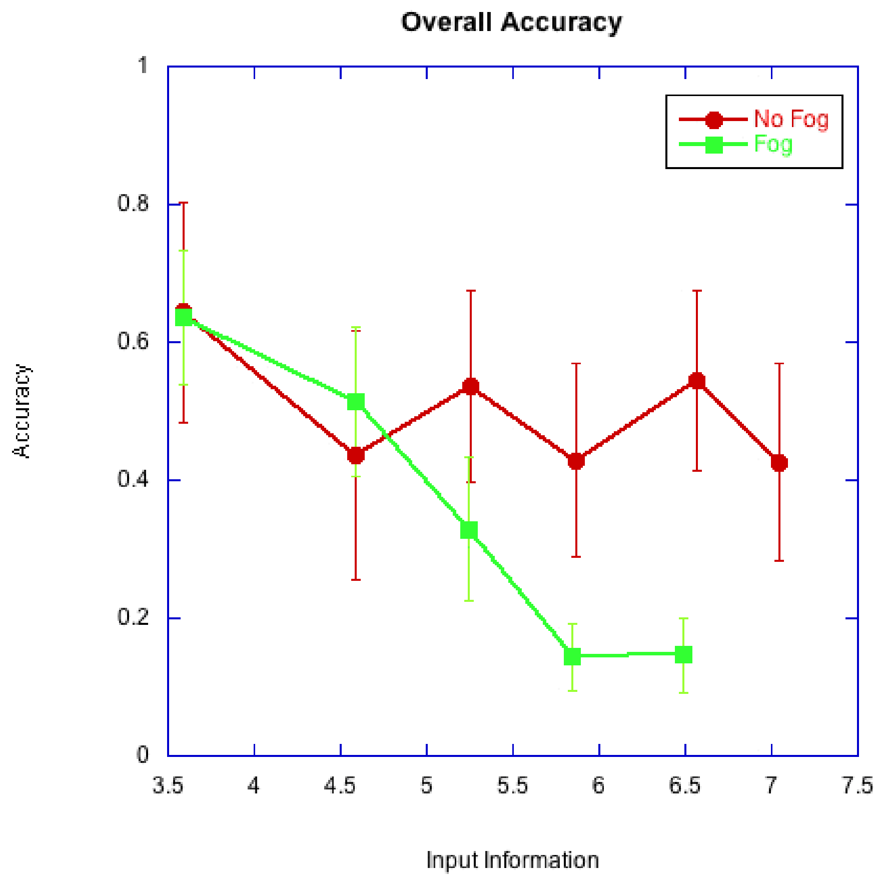


Figure 12: Plot shows the accuracy of participant's averaged over each environment. The green line shows the data for participants in the fog condition and the red line shows the accuracy results from the no-fog condition.

The accuracy results show that for the participants in the no-fog condition, their accuracy stayed constant around 50% irrespective of environment size. This supports the idea that people are using about 50% of the information available in the environment to form a proper internal representation of the environment that is helpful for accurate navigation. In the fog condition the accuracy for the smaller environments is as good as the other participants' performance in the no fog condition. However, we do see a steady fall in the accuracy of responses as the environment size gets bigger. This shows us that for the larger environments people are not only using the immediate cues available to localize and orient them self, distant cues appear to play an important role in accurate navigation. Distant cues are visible from a larger number of views in the environment and can help in localizing oneself in a larger number of states in comparison to proximal cues that are visible from fewer views.

CHAPTER 4. EXPERIMENT TWO

Experiment Two replicates Experiment One but with naïve subjects. The main idea behind Experiment Two was to look at how naïve subjects fare in the same task as in the first experiment. Naïve subjects give us a baseline to compare all the other results against. How do naive participants use the information available in an environment to navigate?

4.1 Procedure

The participants in Experiment Three were eleven students (4 males and 7 females) from the University of Texas at Austin. Three of the participants worked in the lab so they were not paid, whereas all the other participants were paid $10 per hour for their participation. We also had six students (1 male and 5 females) from the University of Texas at Austin participate in the fog condition of the study. They were also paid $10 per hour for their participation. The same environments were used as in the first experiment and the participants ran through the environments in the same order as shown in Table 1. The only difference between the methods used in the two experiments was that in Experiment Two subjects started in the *experimental phase* as described in section 3.1.3.

Experiment Two participants did not have to go through the *training phase* or the *testing phase.*

4.2 Results

We calculated the mutual information for each participant.

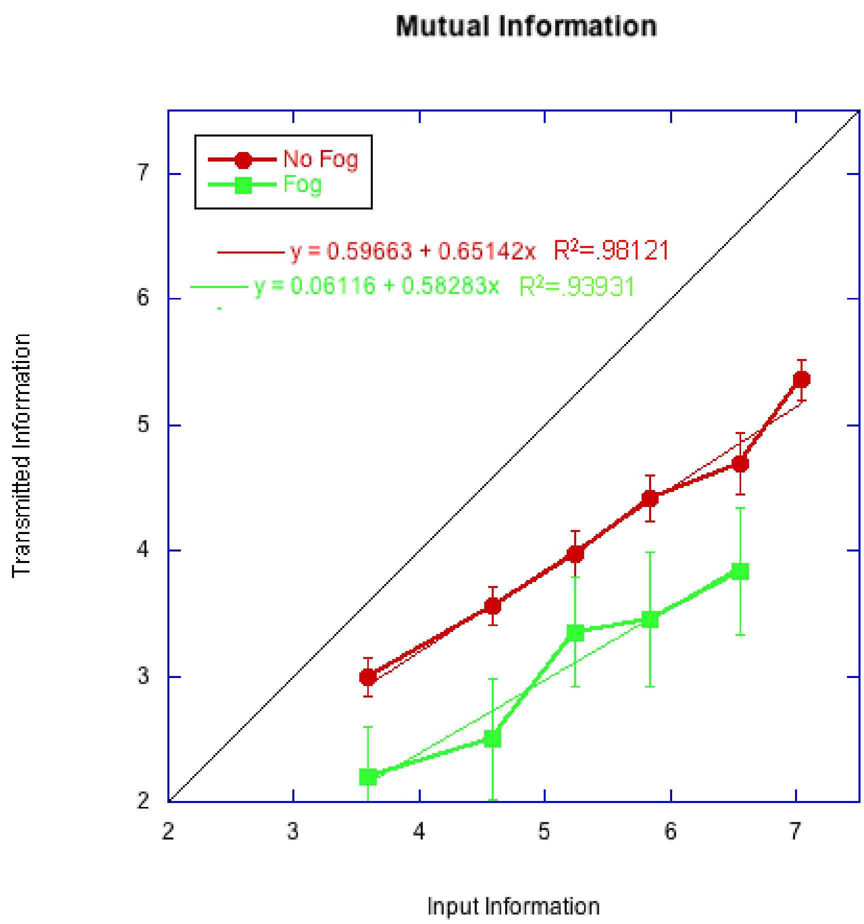


Figure 13: Experiment 3, Mutual Information: for each subject in the six environments. The green lines show the performance in environments with fog whereas the red line is the performance in environments without fog. Linear fit lines (red = no fog, green = fog) are also shown in this plot.

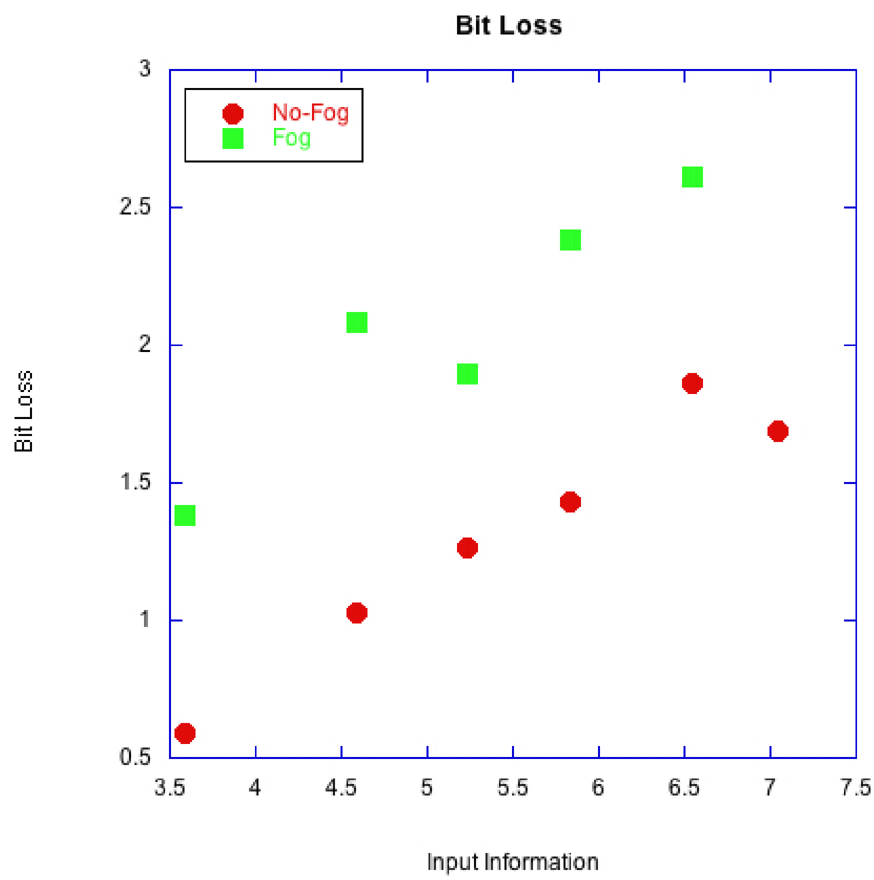


Figure 14: Experiment 2, overall bits lost in each environment. The green squares show bit loss in the fog condition whereas the red circles show the bits lost in the no-fog condition.

Figure 13 shows the mutual information for the participants in each environment. Figure 14 shows the bits lost for each environment. Unlike in Experiment One, naïve participants loose more bits of information in the fog condition versus the no-fog condition.

The results show an imperfect channel with a percentage of the information from the real environment not being transferred into the *Cognitive Spatial Representation* regardless of the size of the environment being tested. The mutual information results are similar to the ones seen in Experiment One. Again we find no channel capacity for up to 7.04 bits or 132 states in either condition of no-fog or fog (100 states or 6.6 bits) but we do see a constant loss of information. Linear fit lines are the best fit for the fog and no-fog data as seen in Figure 13.

The accuracy results as shown in Figure 15 are quite different from the results seen in Experiment One. The chance line (blue) shows how performance would be if participants were only using the structural information from the shown view to pick their response. Naïve subjects in the no-fog condition as seen in the red line have a steadily decreasing rate of accuracy. Their accuracy in the smallest environment is better than the results from Experiment One which is surprising. Participants in the no-fog condition are performing better then performance relative to chance if they were only using the structural information to try and problem solve the solution (blue line in Figure 15). As the environment size increases their performance comes closer to chance. Structural cues from the views are being used by the participants in the no-fog condition. Subjects in the fog condition have very low accuracy (around 10% or lower), irrespective of the size of the environment. In the fog condition performance is below chance as the structural information is removed by the fog. The naïve subjects appear to be unable to form a proper internal representation of the environment that they can use for accurate navigation.

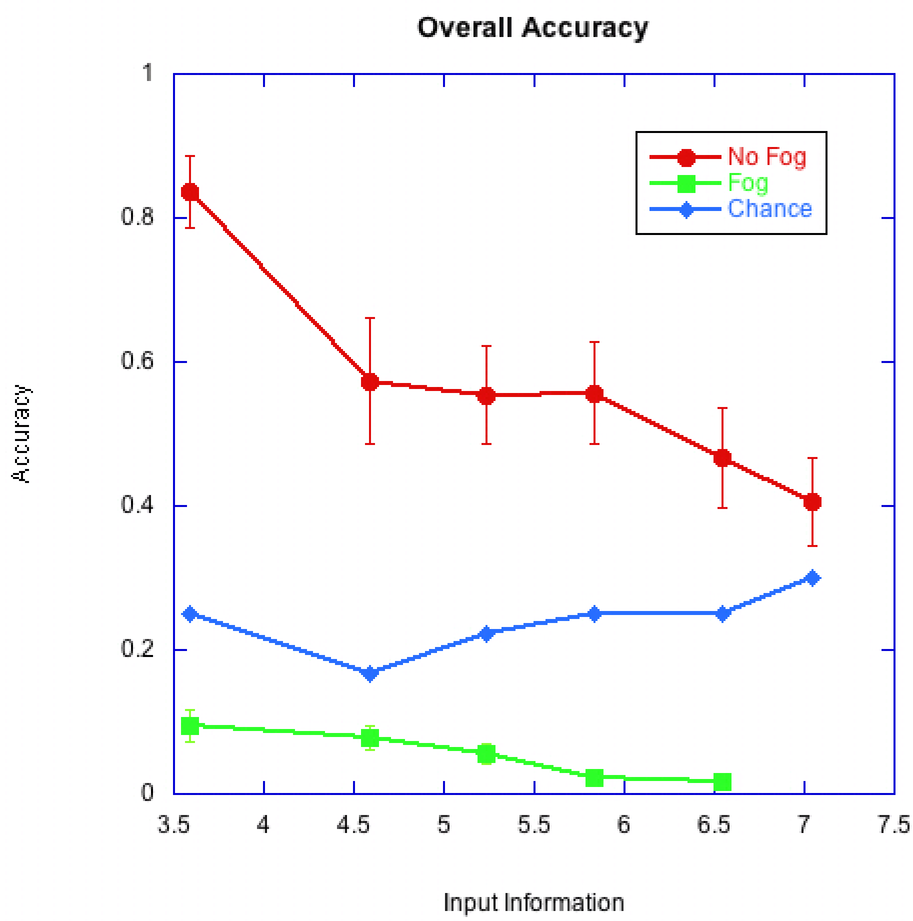


Figure 15: Plot shows the accuracy of participant's averaged over each environment. The green line shows the data for the participants in the fog condition, the red line shows the accuracy results from the no-fog condition and the blue line shows performance at chance that is if only the structural information from a view was being taken into consideration when marking a response on the map.

Data was also analyzed looking only at the first 1/3rd of the trials and the last 1/3rd trials to see if there was any learning taking place during the *experimental phase* itself. Figure 16 shows the mutual information results for the first third (purple line) versus the last third of the trials (red line) and Figure 17 shows the accuracy results for the first third (purple line) versus the last third of trials (red lines). The figures show that there is learning occurring during the *experimental phase* as the performance of the participants improves in the last 1/3 of trials in the no-fog condition.

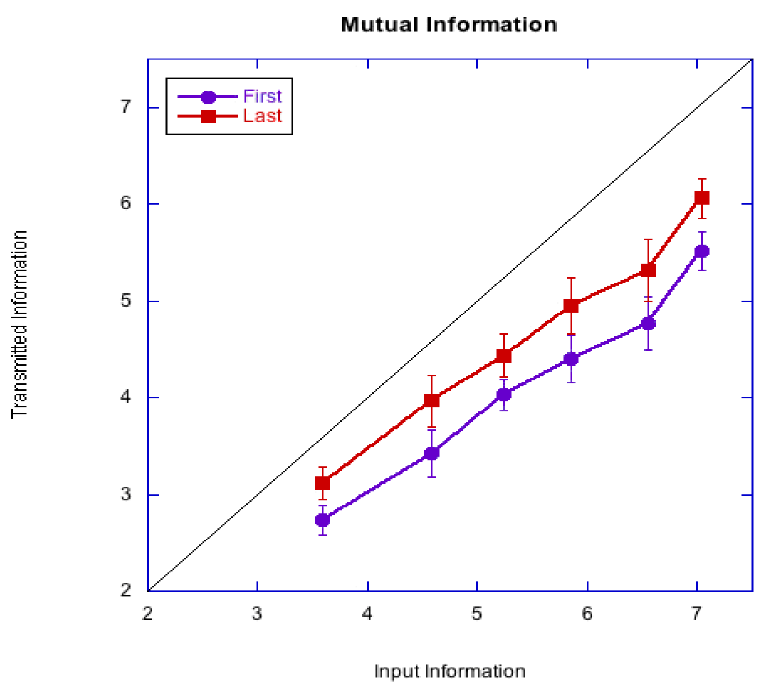


Figure 16: The purple line shows the mutual information for the first 3 trials for each tested state for the participants in the no-fog condition and the red line shows the mutual information for the last 3 trials.

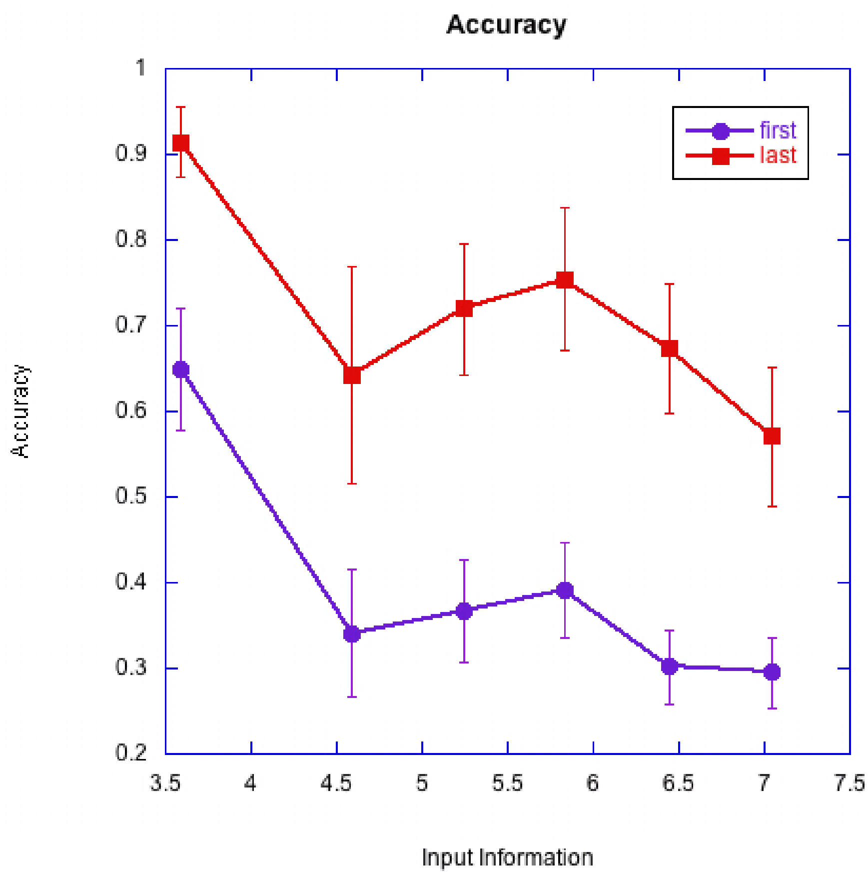


Figure 17: The purple line shows the data from the first 3 trials for each tested state for the participants in the no-fog condition and the red line shows the accuracy results from the last 3 trials.

4.2.1 Results Comparison: Experiment One & Two.

We also calculated the normalized distance error for the no-fog data from Experiment One and Two as seen in Figure 18. Normalized distance error measures the corridor length distance (shortest route) of the inaccurate response from the correct response. When the participant marked the wrong area on the map for a given view, how many corridor lengths away was the incorrectly given response state from the accurate state. The maroon line shows the results of the trained participants from Experiment One and the blue line shows the results from the naïve participants in Experiment Two.

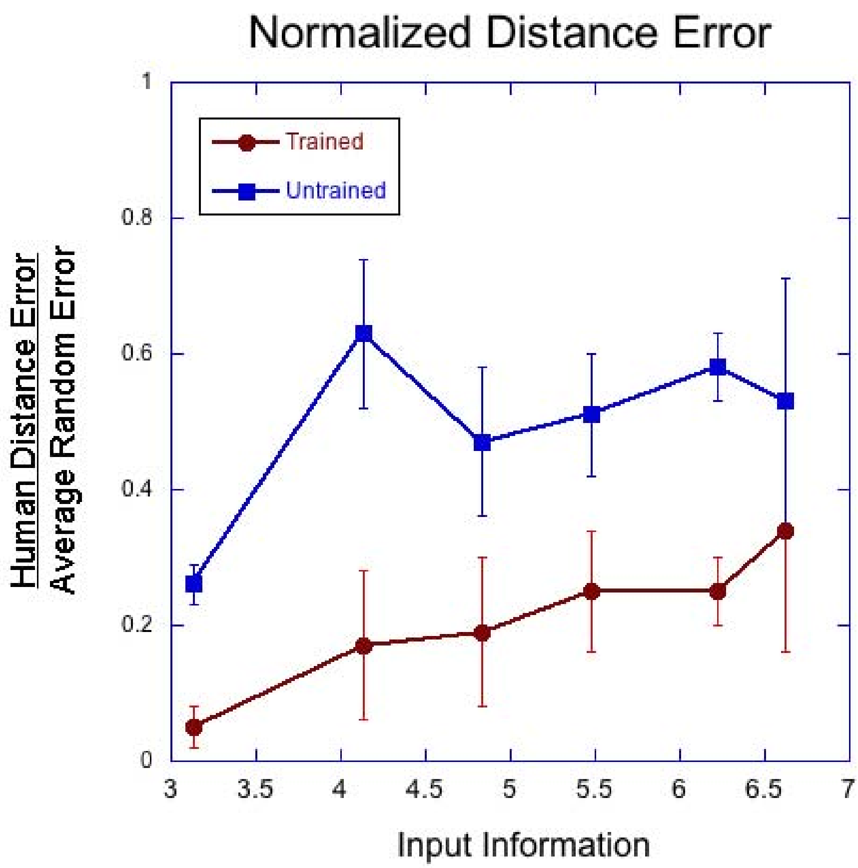


Figure 18: Normalized distance error comparison for trained and naïve participants from Experiment One (blue line) and Experiment Three (maroon line).

In Figure 18, we see a significant difference in distance errors between trained and naïve participants. Trained participant's inaccurate responses were closer to the correct response compared with the naïve subject's results. Even in the two corridor environment the distance errors are larger for naïve subjects although their overall accuracy was better then that of the trained subjects.

Figure 19 (a-x) shows the response errors most commonly made by the participants in Experiment One and Two for each environment. The x-axis shows the percentage of overall correct and incorrect responses. The incorrect responses are further divided according to the type of error made: structural aliasing, orientation errors, structural errors and errors made less than 10% of the times. The structural aliasing, orientation and structural errors are based on the errors made at least 10% or more of the times for each tested view. Each view was tested 10 times in a random order and if the incorrect response for that view was made more then 10% of the time then the errors are further divided according to the type of error. Structural aliasing errors are when the incorrect response has the same structure as the structure of the tested view (for example, being tested on a view of a T-junction and marking another T-junction in the environment in response). Orientation errors look at the incorrect responses where the responded position is correct but the orientation is wrong (for example left-right errors). Structural errors show the percentage of incorrect responses where the structure of the responded position is different from the structure of the tested view (for example being tested on a T-junction view and responding with an L-junction view on the map of the environment). Finally the incorrect responses are further divided into errors made very infrequently, less than 10% of the times. These errors are not further divided as structural aliasing or orientation errors. A large percentage of less than 10% errors show that there was no consistent confusion between a specific tested view and the response given for it. Every time a specific view was tested (each view was tested 10 times in a random order) the incorrect responses were not consistently in the same location but spread over a number of varied locations. Appendix A and B show the detailed error matrix for each environment from Experiment One and Two respectively. All of the participant's responses for each tested state are summed across the environment, showing the overall accuracy confusions of all the participants in a particular environment.

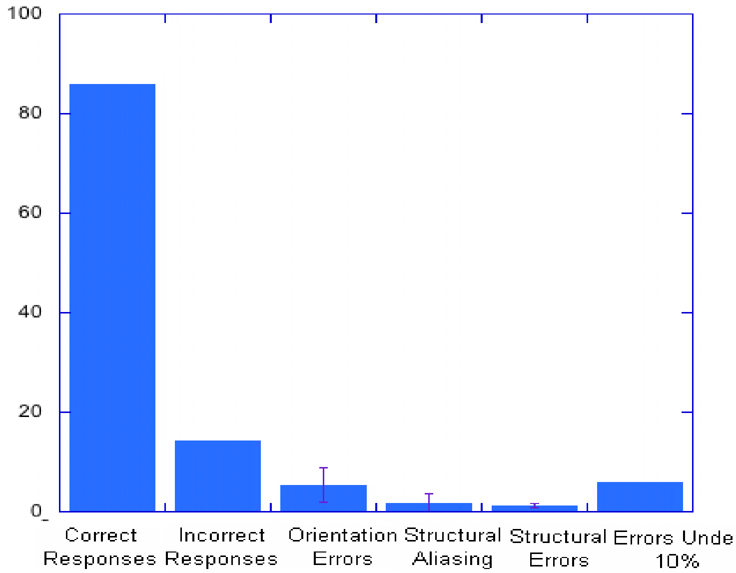


Figure 19A: Env 2 (Exp: 1, Trained), 2 participants, 240 number of overall trials across both subjects

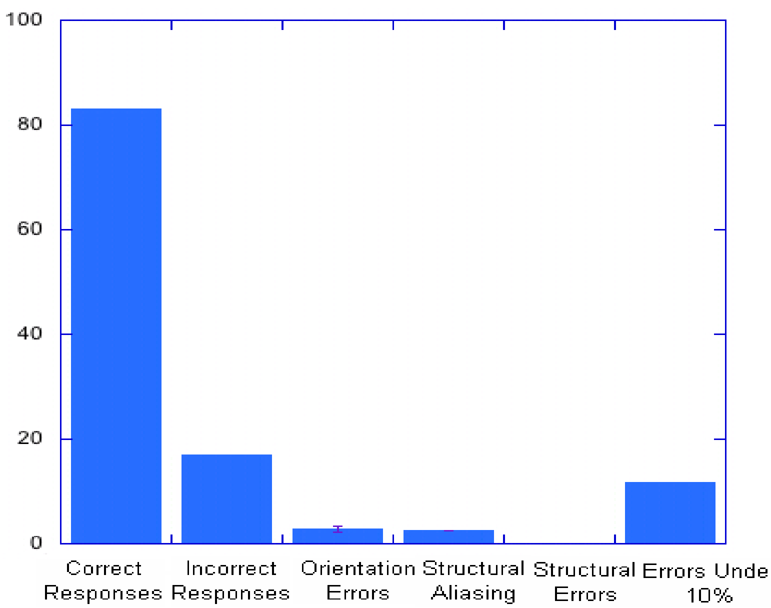


Figure 19B: Env 2 (Exp: 2, Naive), 5 participants, 600 number of overall trials across both subjects

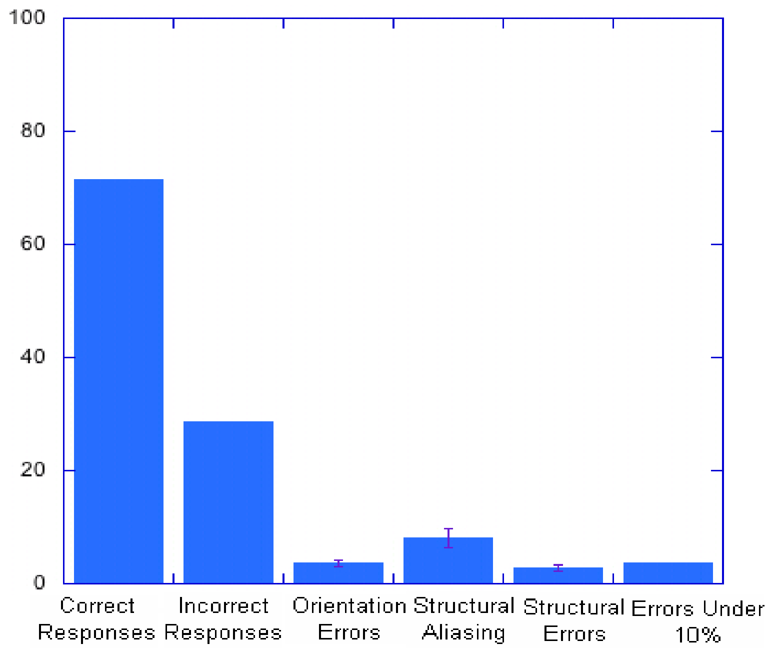


Figure 19C: Env 2b (Exp1, Trained), 3 participants, 360 number of overall trials across all subjects

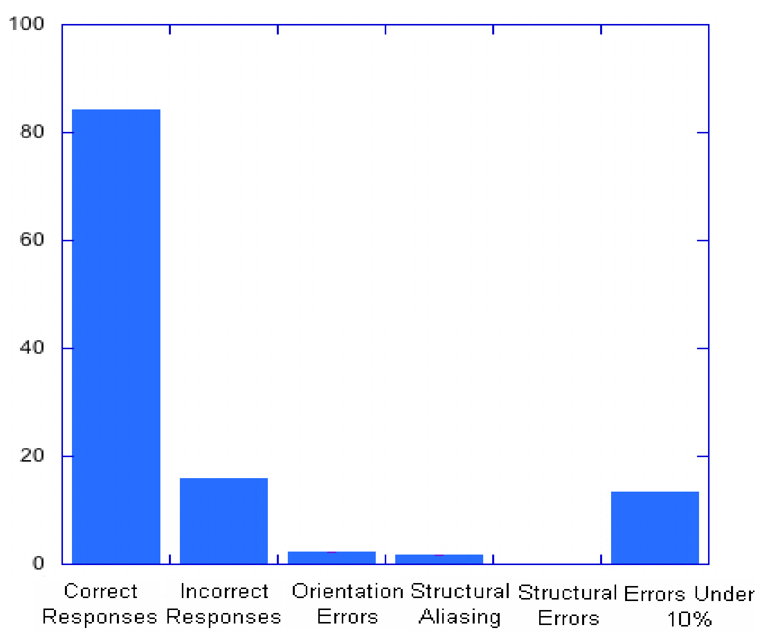


Figure 19D: Env 2b (Exp2, Naïve), 6 participants, 720 number of overall trials across all subjects

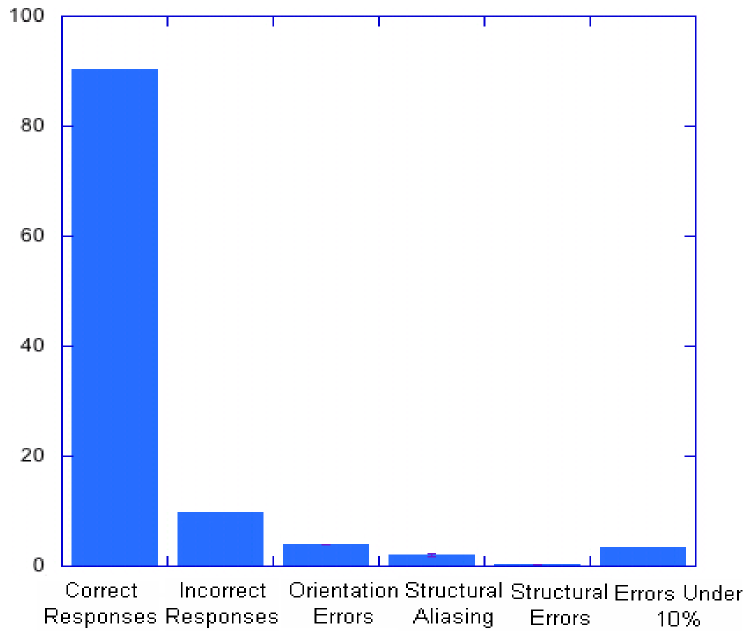


Figure 19E: Env 5 (Exp1, Trained), 2 participants, 480 number of overall trials across all subjects

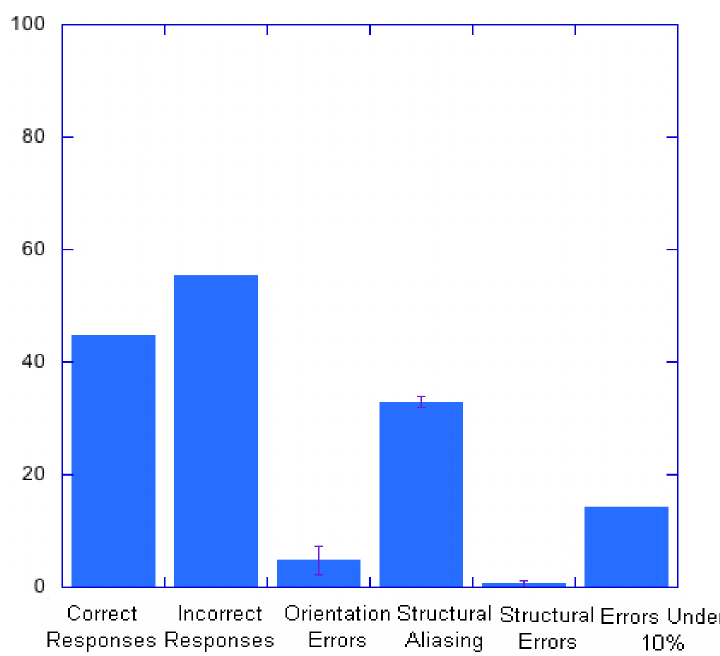


Figure 19F: Env 5 (Exp2, Naive), 7 participants, 1680 number of overall trials across all subjects

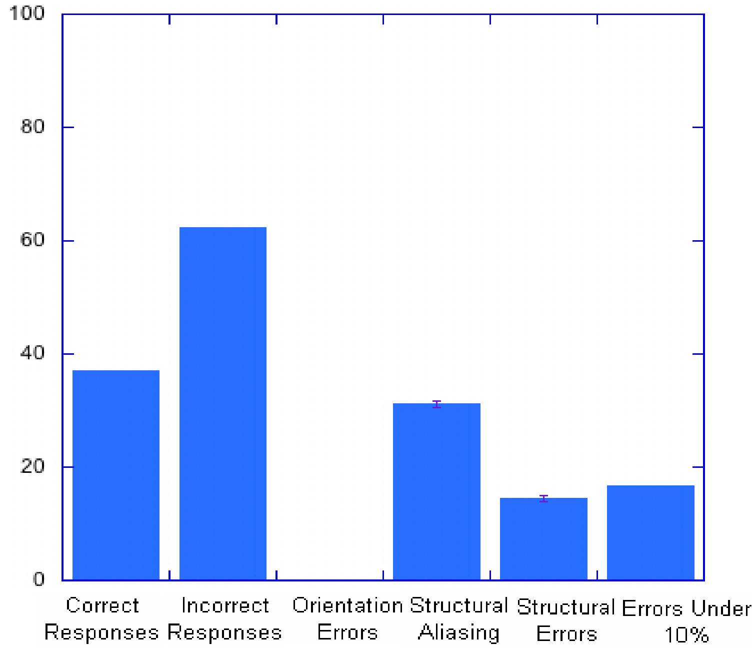


Figure 19G: Env 5b (Exp1, Trained), 2 participants, 480 number of overall trials across all subjects

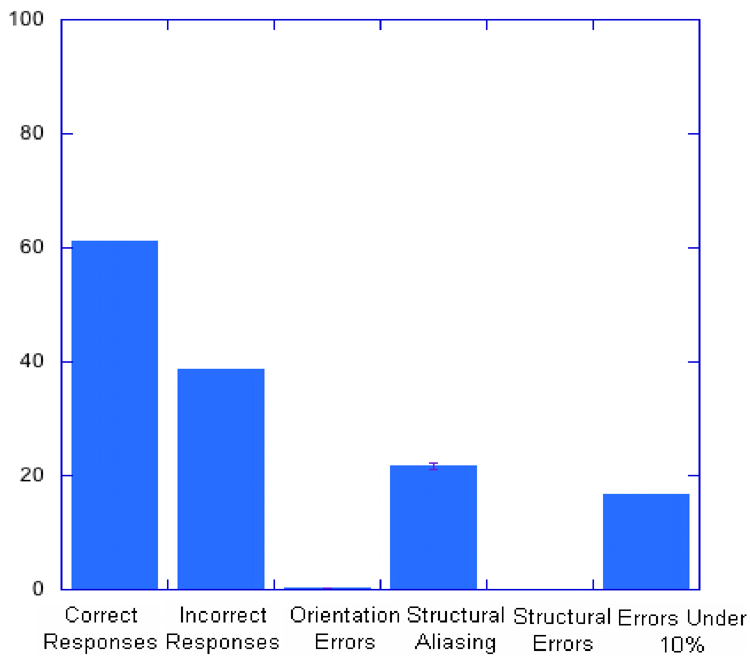


Figure 19H: Env 5b (Exp2, Naive),42 participants, 960 number of overall trials across all subjects

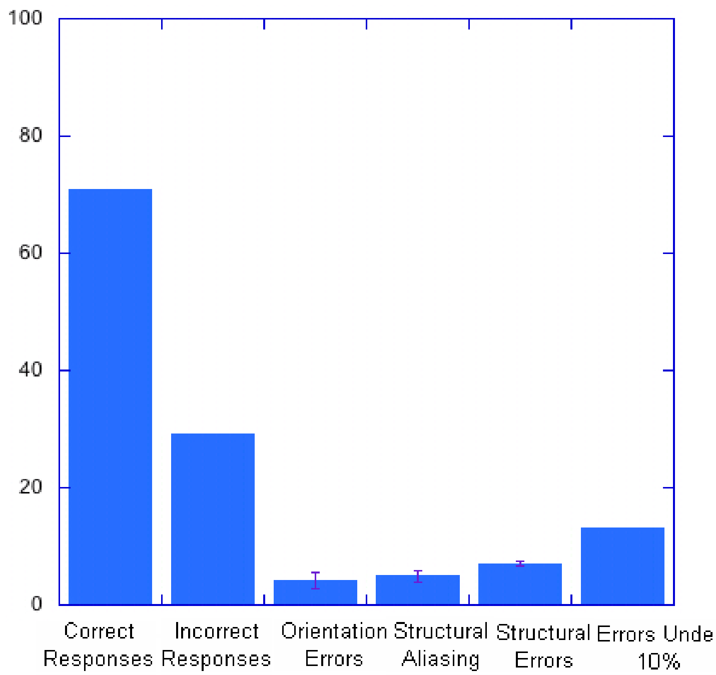


Figure 19I: Env 10 (Exp1, Trained), 3 participants, 1200 number of overall trials across all subjects

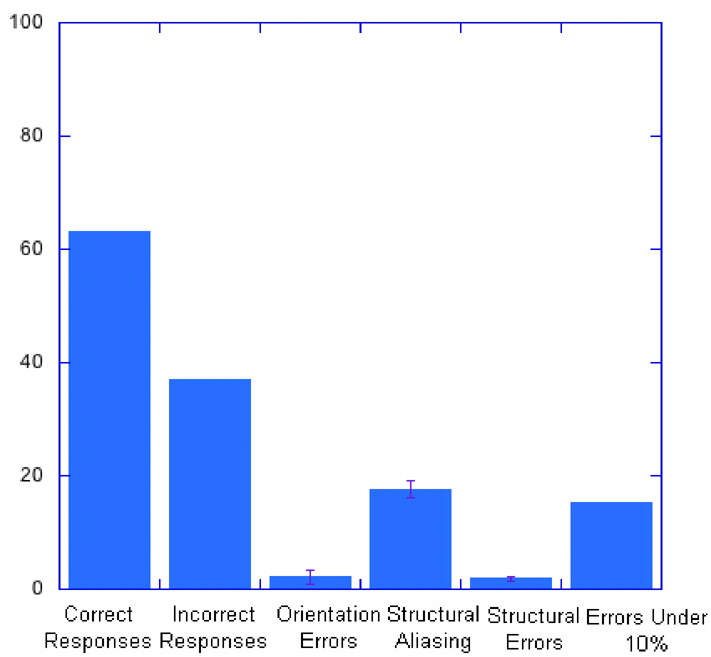


Figure 19J: Env 10 (Exp2, Naive), 5 participants, 2000 number of overall trials across all subjects

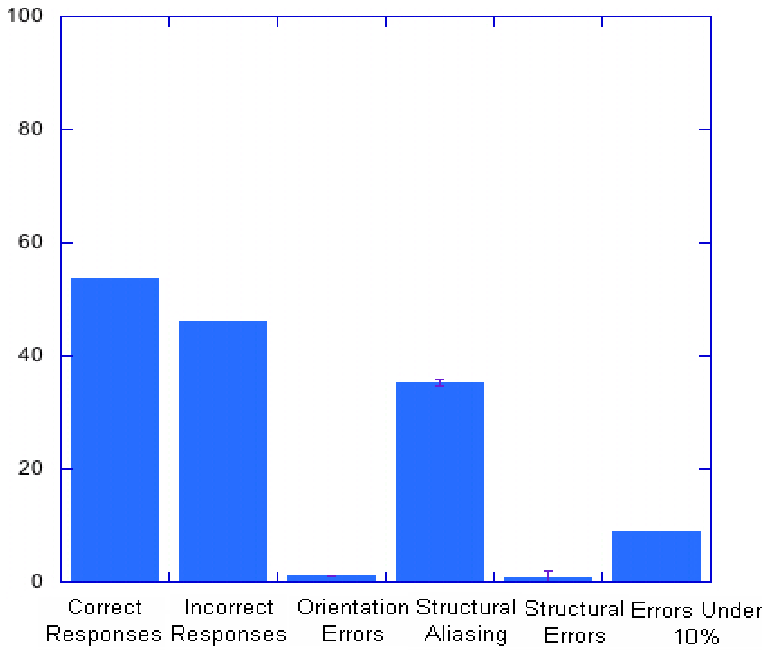


Figure 19K: Env 10b (Exp1, Trained), 2 participants, 720 number of overall trials across all subjects

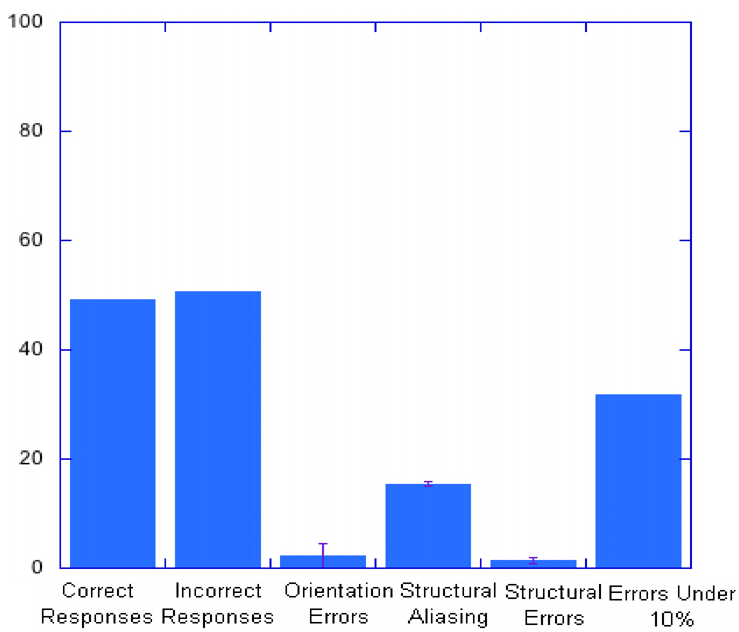


Figure 19L: Env 10b (Exp2, Naive), 6 participants, 2160 number of overall trials across all subjects

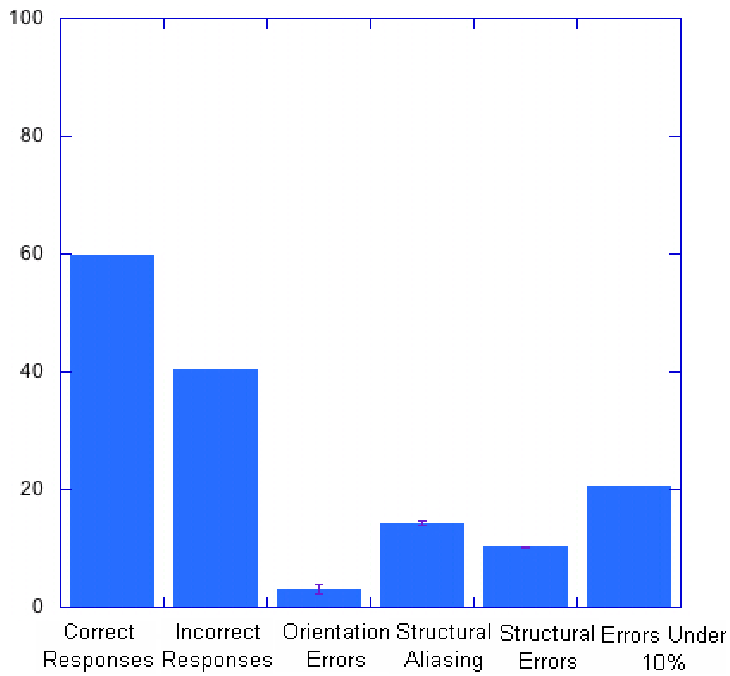


Figure 19M: Env15 (Exp1, Trained), 2 participants, 1120 number of overall trials across all subjects

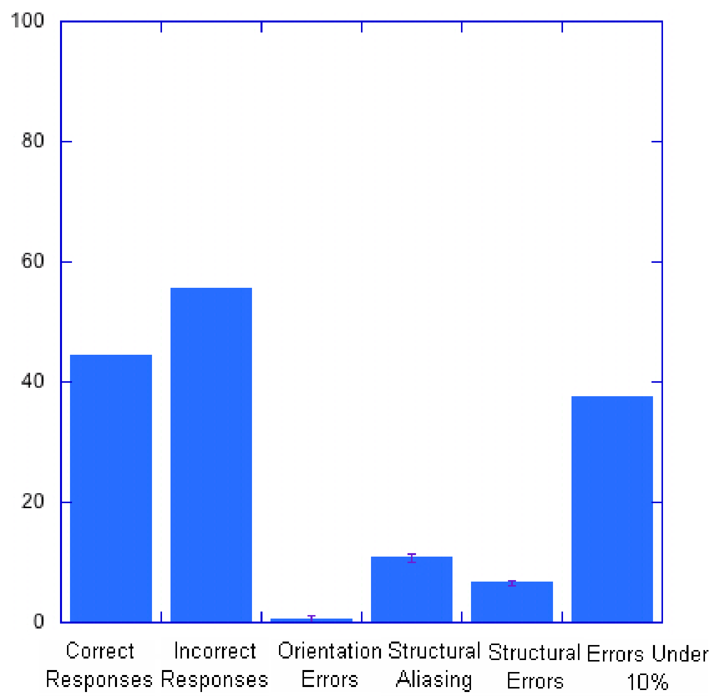


Figure 19N: Env15 (Exp2, Naive), 7 participants, 3920 number of overall trials across all subjects

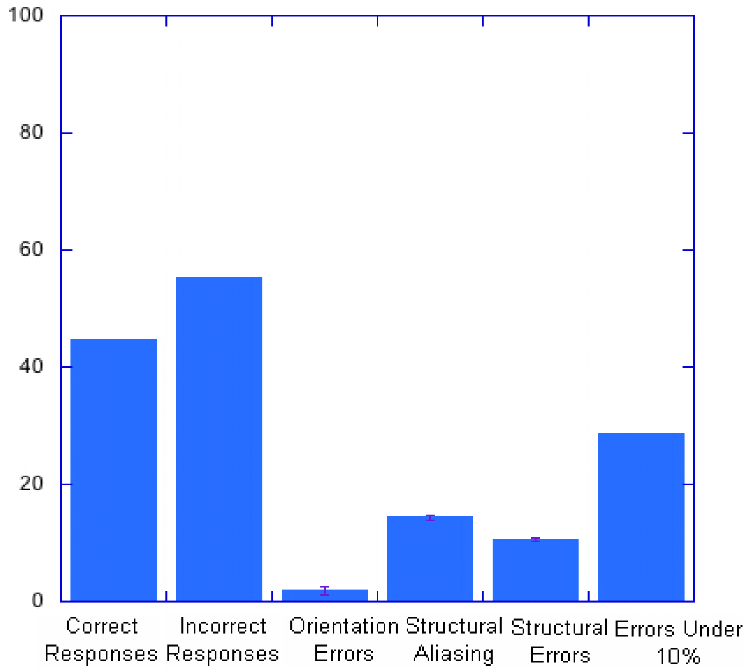


Figure 19O: Env15b (Exp1, Trained), 3 participants, 1800 number of overall trials across all subjects

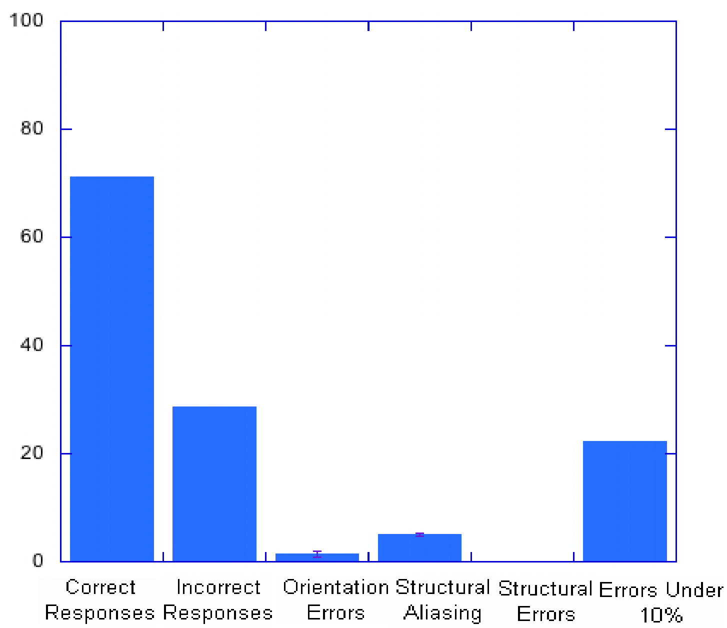


Figure 19P: Env 15b (Exp2, Naive), 4 participants, 2400 number of overall trials across all subjects

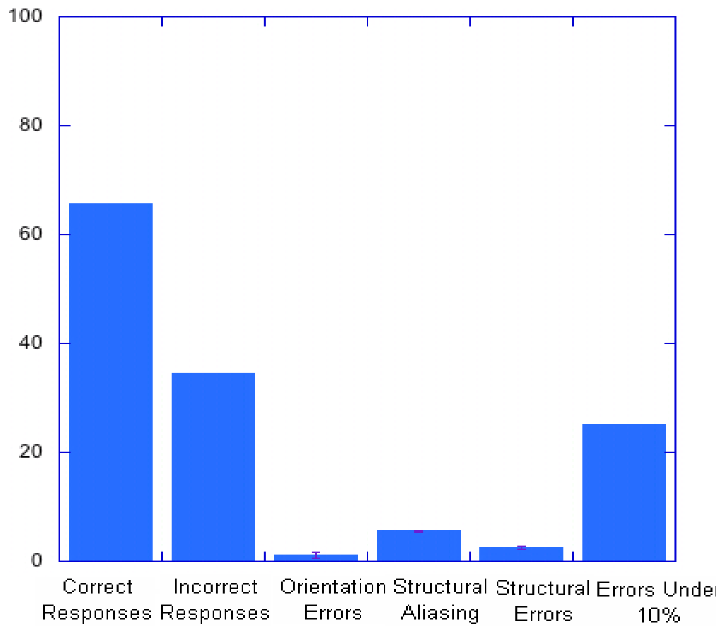


Figure 19Q: Env25 (Exp1, Trained), 3 participants, 3000 number of overall trials across all subjects

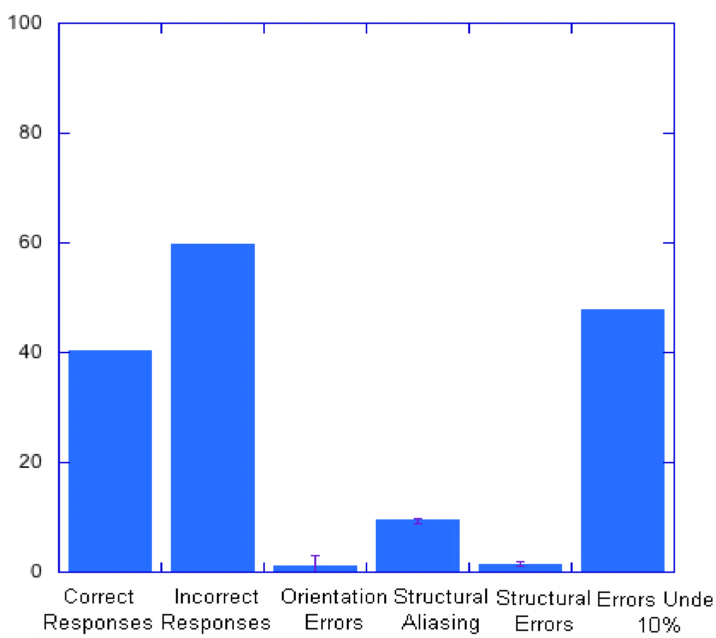


Figure 19R: Env 25 (Exp2, Naive), 6 participants, 6000 number of overall trials across all subjects

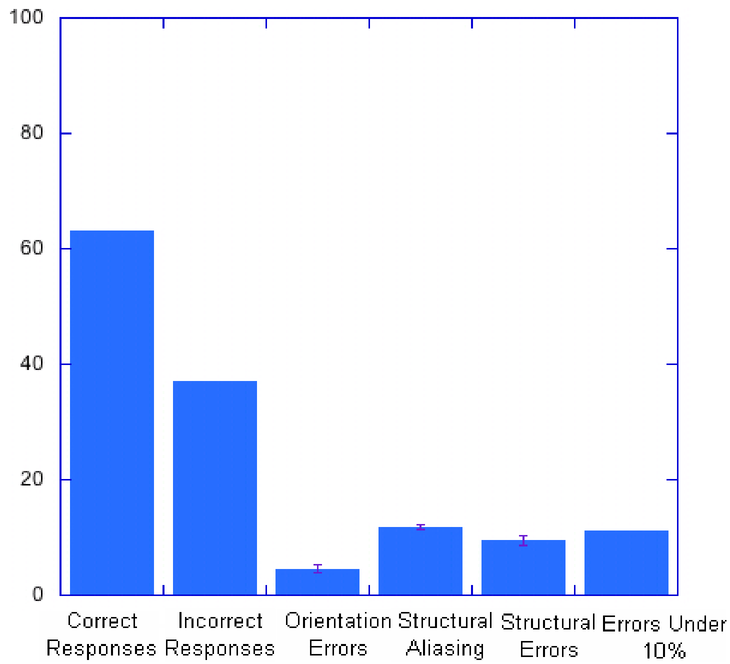


Figure 19S: Env25b (Exp1, Trained), 2 participants, 1760 number of overall trials across all subjects

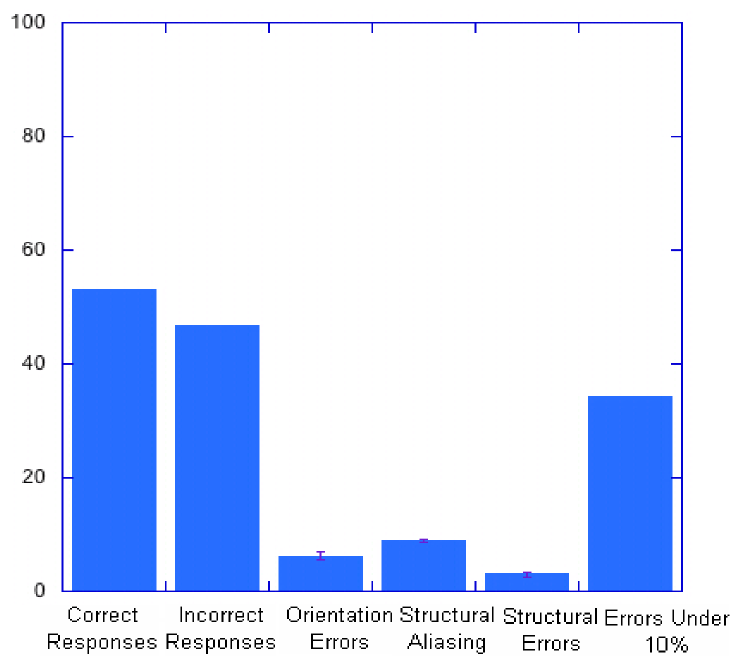


Figure 19T: Env25b (Exp2, Naive), 5 participants, 4400 number of overall trials across all subjects

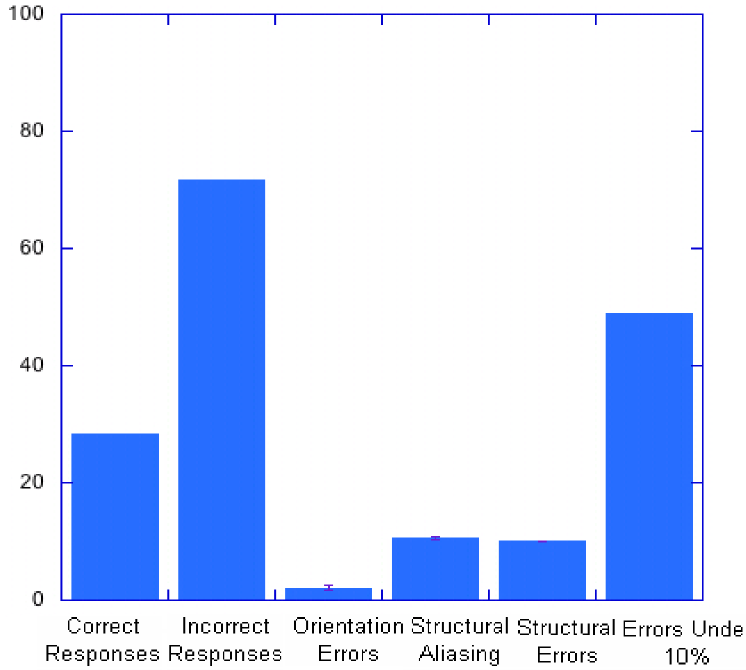


Figure 19U: Env40 (Exp1, Trained), 3 participants, 1980 number of overall trials across all subjects

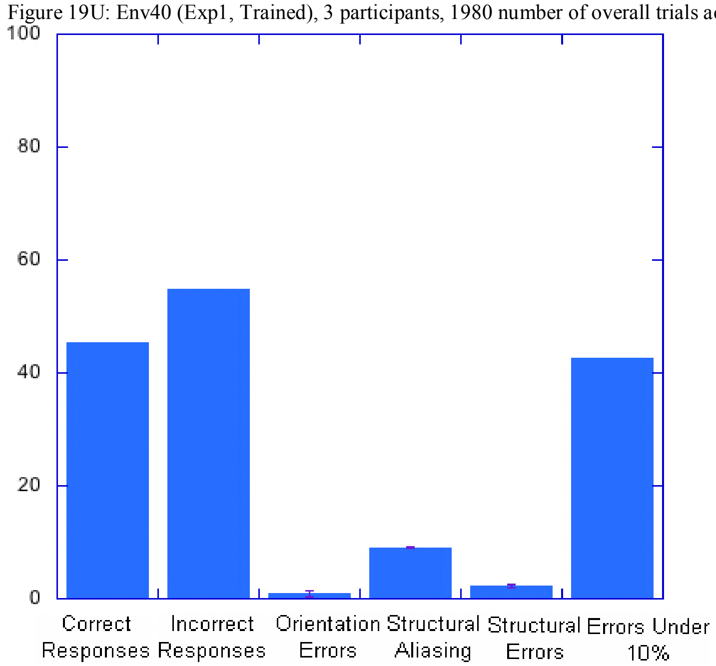


Figure 19V: Env40 (Exp2, Naive), 6 participants, 3960 number of overall trials across all subjects

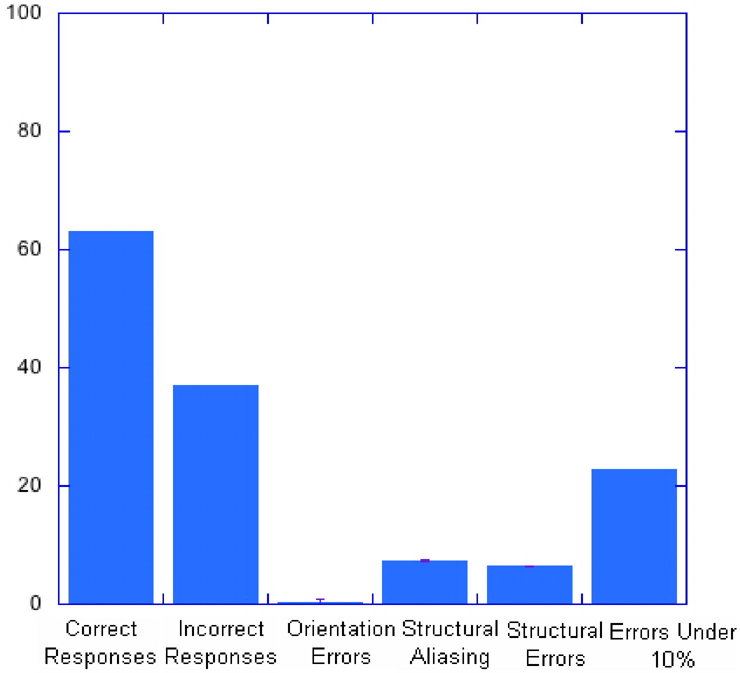


Figure 19W: Env40b (Exp1, Trained), 2 participants, 1320 number of overall trials across all subjects

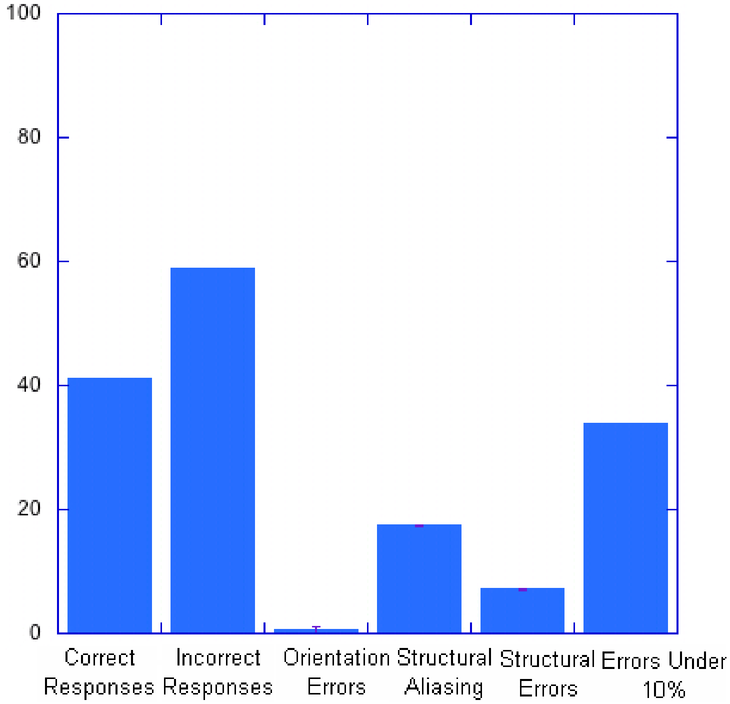


Figure 19X: Env40b (Exp2, Naïve), 5 participants, 3300 number of overall trials across all subjects.

The response pattern for the naïve subjects and trained subjects is quite different. For the smaller environments (2 and 5 corridors) both trained and naïve participants gave more correct responses than incorrect responses. Naïve participants tend to make a larger percentage of 'under 10% errors' compared to trained participants. That shows that when the naïve participants gave an incorrect response for a tested view they tend to mark different positions as their response. They do not consistently mark the same spot for a given view. The results for the trained environment 5B (map of environment 5B can be seen in Appendix A) are surprising as the percent of incorrect responses is higher than the percent of correct responses. This may be due to the structure of the environment as it is perfectly symmetrical. Further analysis of the error results showed that the participants' errors predominantly consisted of structural aliasing errors, showing the impact of the symmetrical nature of the environment on performance. Even the naïve participants have a large percentage of structural aliasing errors for Environment 5b.

Naïve participants start to have a larger percentage of incorrect responses versus correct responses once the environment size starts getting larger than 10 corridors. Trained participants for most of the larger environments have a higher percentage of correct responses. The 'under 10%' errors are a higher percentage of the error results for the naïve participants than the trained participants for almost all the environments. The structural aliasing errors also make a large part of the errors especially for the trained participants across all environment sizes. Structural aliasing errors show that the participants are using the structural information to localize themselves but they do not have enough landmark information to accurately localize themselves in the environment. Orientation errors do not make a large percentage of the error responses for the naïve or trained participants.

Trained participants also tend to make a larger percentage of structural errors than naïve participants. As the naïve participants had no learning experience they are using the structure from the view to pick their responses hence they have a smaller percentage of structural errors. Whereas the trained participants who have had learning experience in the environment have other sources of information available to them from their experience in the environment and are not only using the structure to pick their response. Hence we see a higher percentage of structural errors for the trained participants across all environments. The trained participants are using more than just the structure of the environment to localize themselves.

CHAPTER 5. EXPERIMENT THREE

The main idea behind Experiment Three was the same as in Experiment One and Two, but instead of using virtual worlds we wanted to look at what information was being transferred in the real world context. The question arises whether the desktop navigation results will translate to more realistic conditions. Our lab has conducted studies in which we compare navigation performance in three different conditions: key-press condition (as in our virtual environment studies), joystick condition allowing continuous movement rather than the quantized movements and immersive condition where the participant's movements in a virtual reality arena are recorded. In each condition the participant's efficiency in reaching the goal state was measured. Results did not show any significant difference in the performance across all three conditions. These results support the use of desktop environments to study navigation. Virtual environments are advantageous as they allow the experimenter greater control in the topology, placement of landmarks, size, etc. of the environment (Stankiewicz, Legge, Mansfield, & Schlicht, 2006).

For testing purposes we picked the West Mall area in the University of Texas at Austin campus. We divided the area into three different sized environments. The smallest environment consisted of 8 testable states and the largest environment consisted of 128 testable states as shown in Figure 20. Our main experimental manipulation was varying the number of views to be tested.

Fig 20: Campus map with 128 testable states.

Pictures were taken at each test location. At each location four pictures were taken in all four ordinal directions (north, south, east and west). Hence there were four pictures taken at each position, showing the view from all four directions. Example of the pictures used can be seen in Figure 21.

Fig 21: Campus view

5.1 Procedure

Sixty undergraduate students (Thirty six females and twenty four males) from the University of Texas at Austin were given class credit to participate in the current experiment. The participants were divided into three groups of twenty people each, in a random order.

When a participant came in for the study, they were given a questionnaire to complete. The questionnaire asked about the student's navigational experience on campus and their knowledge about the tested area. Some questions used were:

1) How many years have you been at UT?

2) Name ten buildings on campus that you spend most of your time in (academically or non- academically). Please rank the buildings according to the amount of time you spend there (one being the most visited and ten being the least visited)?

3) Do you drive, walk, ride the shuttles or ride a bike to class?

There was no exploration phase in this study as the subjects already had ample experience in the environment as validated by the questionnaire responses.

5.1.1 Experimental Phase

The participants used a desktop computer for the experiment. During the *experimental phase* the subjects were shown a picture of a specific view from the environment as shown in Figure 22a. Once the subject was ready, they hit the space bar and were shown a map of the environment as shown in Figure 22b. The subjects were not shown the map of the environment until they started the *experimental phase*. The main streets were named on the map along with a captioned picture of the Main Building. These were the only cues available to the participants to help them localize themselves to the specific region on campus.


Figure 22a: Campus View

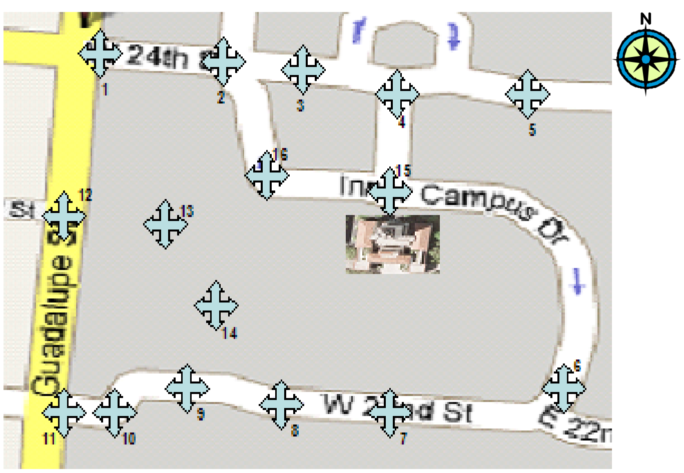


Figure 22b: Campus Map

Each tested state was given a unique number on the campus map, and a compass was also present. The subjects had to write down the location and orientation of where the view was generated from. Subject's answer would consist of the unique number of the response state and the orientation (north, south, east or west). The participants were only allowed to state one spot as their response. Each state was tested three times in a random order.

To counter any picture order effects, there were two versions (A and B) of each environment. The pictures remained the same but the order in which they were presented was changed. A participant was randomly chosen to run in version A or B.

5.2 Results

Mutual Information was calculated for each subject.

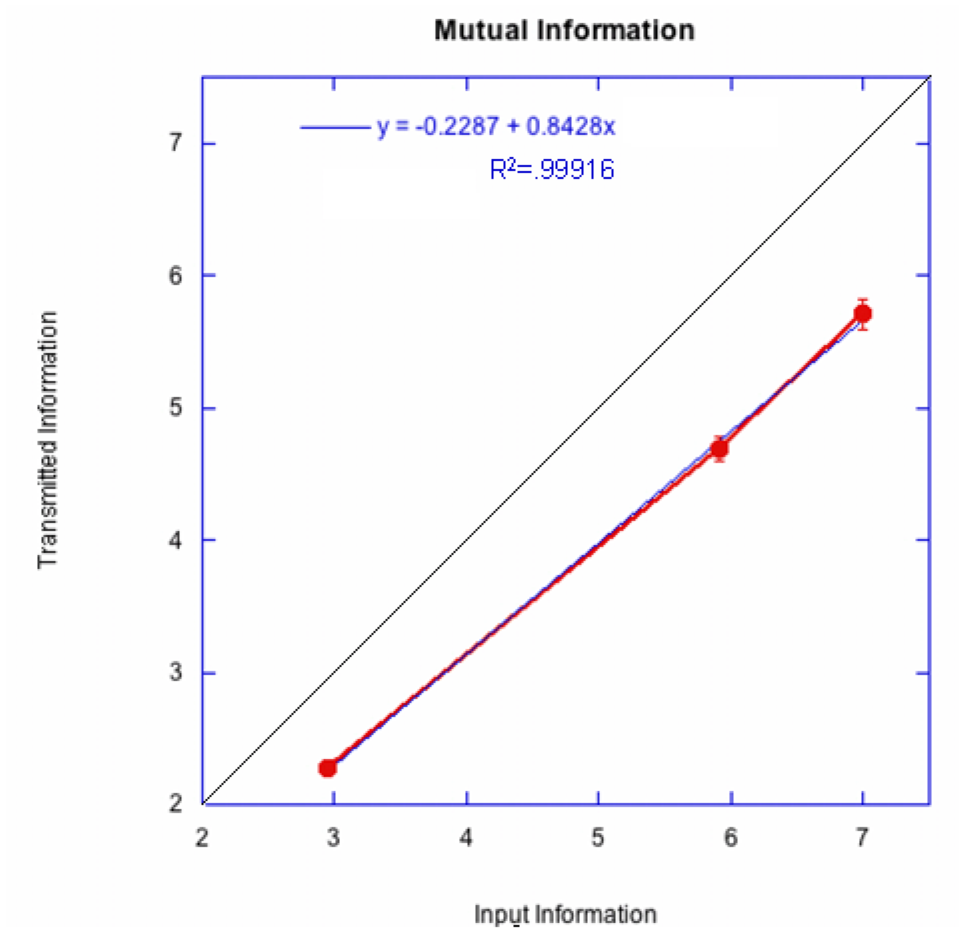


Figure 23: Mutual Information (red line) calculated for each environment. Best fit is a linear fit line as seen by the blue line.

Figure 23 shows the mutual information for the participants in each environment. The results show an imperfect channel with a percentage of the information not being transferred to the *Cognitive Spatial Representation* regardless of the environment size being tested. The mutual information results are similar to the ones seen in Experiment One and Two. Again, we find no channel capacity for the different sized environment of 6, 64 or 128 testable states but we do see a constant loss of information. The best fit line for the date is a linear fit line (blue line) as seen in Figure 23. In realistic conditions too participants do not appear to be using all of the information available in an environment. They are using a percentage of the information available. Transferring a sub set of the information can help overcome the limitations to the amount of information that can be stored in memory.

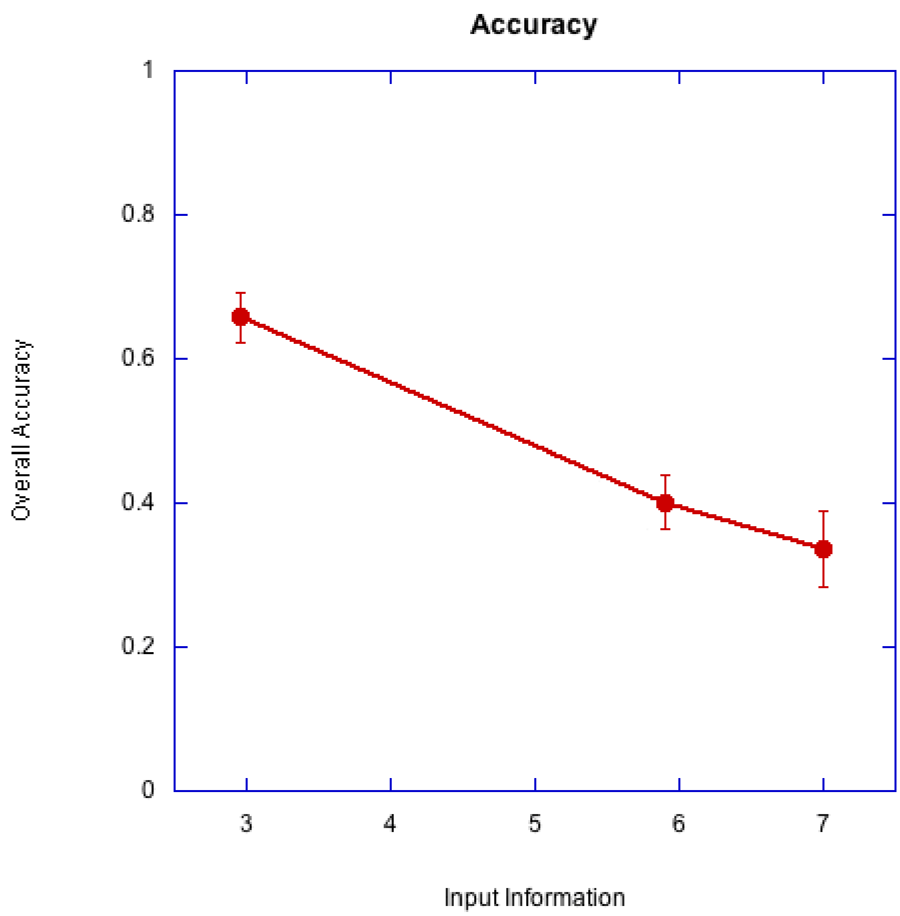


Figure 24: Accuracy for each environment

Figure 24 shows the accuracy results for each environment. Here we see that the accuracy of responses decreased according to the environment size unlike the mutual information results. For the smallest environment accuracy is around 60% which was the same for Experiment One participants. Unlike our earlier results we do not see steady accuracy results of around 50%. In Experiment Three we see a decrease in accuracy as the environment gets bigger from 6 tested states to 64 tested state. Figure 25 shows the accuracy confusions in the data. That is, when the subjects gave an inaccurate response how far were they from the right answer.

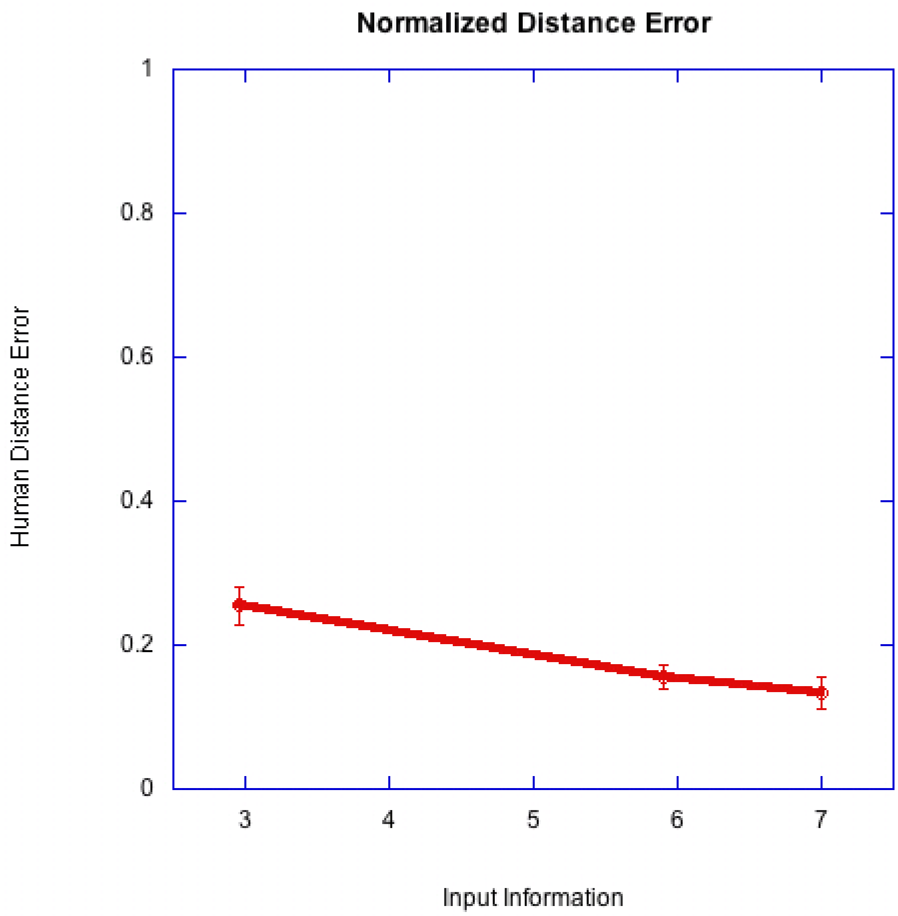


Figure 25: Normalized distance error comparison for participants in Experiment Three.

In Experiment Three as in Experiment One, participant's inaccurate responses were closer to the correct response. Environment size does not greatly affect how far off from the right response the participants were.

As analyzed for Experiment One and Two, Figure 26(a-c) shows the response errors most commonly made by the participants for each environment. As the structure in the real world environment as not as regular as in the virtual world these plots focus on the orientation errors (e.g. left-right errors) and how far off from the right response was the incorrect response (1 hall length or 2 hall lengths). A position marked incorrectly less than 10% of the times by all the participants for a particular tested view is not plotted in these figures. The plots also show the percentage of structural error responses made; that is when the incorrect response was in a completely different area of the environment than the tested view.

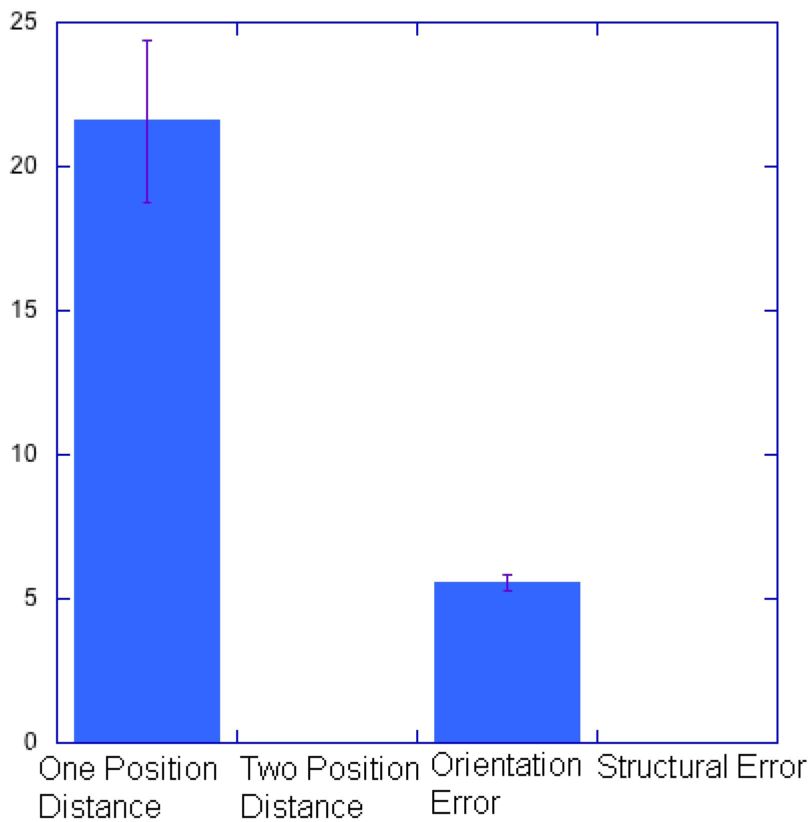


Figure 26A: 8 tested states,576 number of overall tested trials

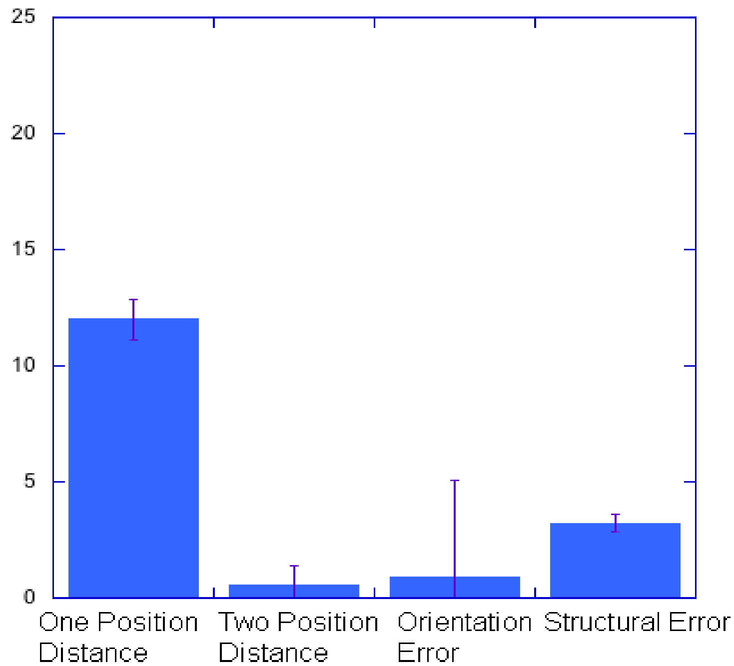


Figure 26B: 64 tested states, 6144 number of overall tested trials

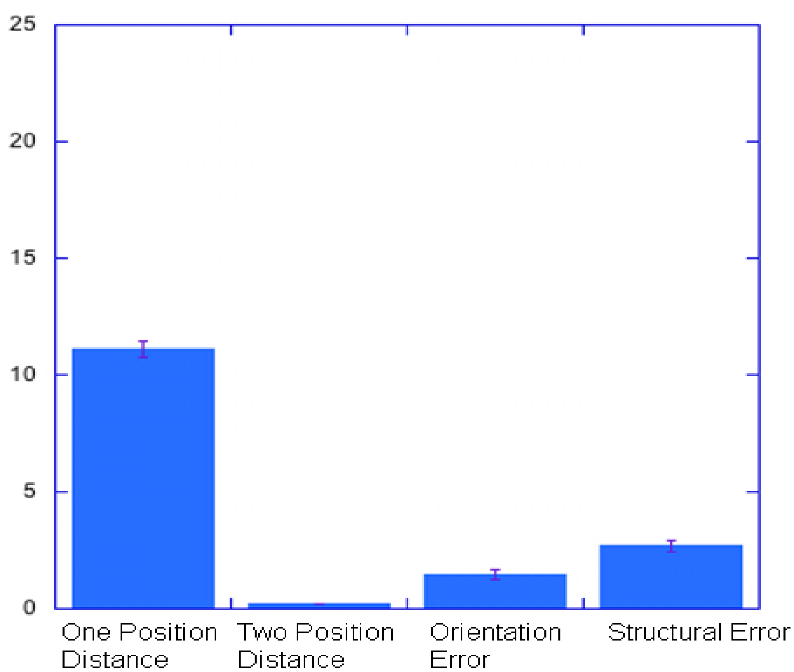


Figure 26C: 128 tested states, 8832 number of overall tested trials

Plots 26 (a-c) show that the highest percentage of errors made across all the environments were incorrect responses that were one position away from the correct responses. Orientation errors only do not make a large percentage of the errors.

CHAPTER 6. KNOWLEDGE ACQUISTION

The results from the first three studies have consistently shown that there is a loss of information, regardless of the complexity of the environment. That is, people appear to be using about 50% of the information available in well-learnt large-scale spaces. Our next step is to look at what kind of information makes up the 50% that is being stored in memory. Encoded information may be different from what is perceptually available to people in a large-scale environment. What specific information is retained and recognized within large-scale spaces (buildings or cities) that play an important role in way-finding and localization abilities.

According to Siegel and White (1975), acquisition of spatial knowledge begins with the learning of landmarks in a new environment. Eventually the landmarks become a part of route knowledge that leads up to the formation of survey knowledge of the new environment. Subsequent research argues more for a parallel acquisition of landmark and route knowledge rather then a step by step process (Moar & Carleton, 1982). Children as young as 8 years also show a preference for learning landmarks that are easily visible on a route that they are using (Cornell, Heth, & Alberts, 1994; Cornell, Heth, & Rowat, 1992). Research by Mallot and Gillner (1998) showed that people perform poorly in navigation tasks when landmark positions are changed. This demonstrated the importance of landmarks for the *view-graph* approach as people's performance is affected detrimentally when the changed landmark position give conflicting information as to the action (e.g. go straight vs. turn) to be taken (Mallot & Gillner 2000; Scholkopf & Mallot, 1995).

S*patial semantic hierarchy* (Kuipers 2000, 2001; Kuipers & Byun, 1991) is a model that depends on multiple interacting representations of a large scale space being used for knowledge acquisition in a large scale space. The type of spatial representation used depends on the navigational task that has to be completed. This model represents space at five distinct levels. The sensory level focuses on the information acquired through sensory (e.g. vision) interaction with the environment during exploration. The control level describes the environment in terms of local control laws (e.g. move forward, turn left) that are activated if the environment has appropriate conditions based on the local geometry of the environment and is terminated once completed. Next is the causal level where schemas and routines can be formed by using the information acquired from the sensory image (including landmarks) when the agent is at a particular position and control levels. The topological level is where the topological map of the views and actions observed at the causal level is formed. The final stage is the metrical level and it is not essential for successful navigation. The metrical level is where one “global geometric map” is formed by combining all the topological maps for the environment (Kuipers, 2000). The *spatial semantic hierarchy* allows for navigation under conditions of uncertainty by allowing for knowledge acquisition from multiple knowledge sources for problem-solving.

As seen in previous research for most theories on spatial learning, gaining landmark knowledge is an important factor for success in navigation. A *landmark* can be defined in many different ways. It can be seen as a reference point in an environment (Lynch, 1960) or distinctive visual features in an environment that can help identify a particular location (Siegel & White, 1975).May, Ross, and Byer (2005), in their research on improving navigation systems found specific characteristics that defined a good landmark. The characteristics were object permanence, visibility, usefulness of the location where the landmark is found, ease in describing the landmark and the degree of interaction that an individual has with a particular landmark while exploring an environment.

Others have divided landmarks into object landmarks and structural landmarks. Object landmarks would be the visual objects present in an environment that are not part of the environments structure. In the current research object landmarks are the pictures placed throughout the environment. Structural landmarks are the geometrical visual cues found in the environment (e.g. T-junction, L-junction, dead-end) (Stankiewicz & Kalia, 2007). In this paper landmarks are defined as salient and easily perceivable features present at specific locations in the environment (Stankiewicz & Kalia, 2007). Research has shown that the presence of object landmarks in an environment helps in navigation and orientation performance (e.g. Abu Ghazzeh, 1996; Darken & Sibert, 1996; Vinson 1999). May et al. (2005), found that in navigational instructions, when good landmarks (as specified in landmark characteristics above) were used they were the primary method used for localization.

However not all landmarks are used equally. Ruddle et al. (1997), showed that participants used highly informative object landmarks (3-D models of everyday objects e.g. cup, fork) when they were present. These landmarks were meaningful as each object landmark in the environment was unique. On the other hand, participants were seen to have difficulty in using less informative and meaningless landmarks (colored abstract paintings) as aids for successful navigation. Object landmarks have to compete with one another to be stored in memory. Once rats have learned to navigate towards a defined goal with reference to specified landmarks they are slow to use new landmarks when they are added to the environment (Rodrigo, Chamizo, McLaren & Mackintosh, 1997; Sanchez-Moreno, Rodrigo, Chamizo & Mackintosh, 1999). This difficulty in replacing already stored object landmarks with new object landmarks shows that not all spatial information is transferred to the internal representation.

Landmarks can act as the strategic focal positions at the beginning and ending of a path and they can also be used as intermediate course-maintaining features. Current research investigates whether the loss of information is strategic (e.g. at non-decision points) or whether there is simply a generalized forgetting function. Aginsky, Harris, Rensink & Beusmans (1997), showed that landmarks at decision points (areas where decisions have to be made about changing direction) are better remembered in comparison to other states. We want to know where the information loss is coming from. Stankiewicz and Kalia's (2007), research results showed that increasing the information content of object landmarks improved participant's retention of objects landmarks. In the current experiments the object landmarks in the virtual environments are highly informative as none of the pictures are repeated in the same environment. Each picture (object landmark) specifies a unique state in the environment. Aim for the study was to determine whether the information that is stored and can be recalled is simply a random sample or has been strategically selected. Previous work has shown that there is a limit to the amount of information being used, leading to the question of what information (focusing on object landmarks in the current work) is being lost in the formation of a *Cognitive Spatial Representation.* Are all of the landmarks stored in memory and easily useable as cues when needed or are there specific landmarks that are learned whereas others are not stored in memory. Ruddle et al. (1997) showed that the form of a landmark affects how it is used, whereas Stankiewicz and Kalia (2007) showed that the information content of a landmark decides if it will be stored. How do people pick the specific landmarks to be used is it dependent on the position of the specific landmark, proximity of the landmark to other landmarks or the distance of the landmark from the viewer (distal cues or immediate cues)?

CHAPTER 7. EXPERIMENT FOUR (A)

Experiment Four investigates whether gaze patterns can reveal what object landmark information is being encoded versus lost in the formation of the *Cognitive Spatial Representation.* During the *experimental phase* half of the landmarks were removed depending on the amount of time the participant spent looking at them during exploration. If the participant is only acquiring some of the object landmark information then removing highly viewed landmarks should have a significant effect on performance but removing seldom viewed landmarks should not affect performance as they are not being used to form the *Cognitive Spatial Representation*. The participants are also tested with all the landmarks present. In this condition performance should be the same as when only the highly viewed landmarks are present, participants would just ignore the seldom viewed landmarks and only use the object landmarks that they transferred to their *Cognitive Spatial Representation.* If participants are learning all of the landmarks in an environment then the removal of some of the object landmarks should not be a problem as the participants should be able to use any of the cues that are present. A preference for any of the object landmarks in the environment will not be seen if they are all being used equally to form the *Cognitive Spatial Representation.*

7.1 Procedure

Fourteen students from the University of Texas at Austin participated in the study. They were paid $10 for their participation. The study had two groups Low-Landmark-Viewing Group and High-Landmark-Viewing Group. The groups differed according to which landmarks were removed during the *experimental phase*. The participants were randomly assigned to one of the two groups, with 7 participants in each group. The study was completed in two sessions on consecutive days. 10 corridor environment from Experiment One and Two as seen in Figure 27 was used for this study. Only one environment was used in this study as capacity limitations on information transfer are not being investigated but the question is how the information in an environment is being used. The 10 corridor environment is a medium sized environment and it contains 40 unique pictures acting as object landmarks.

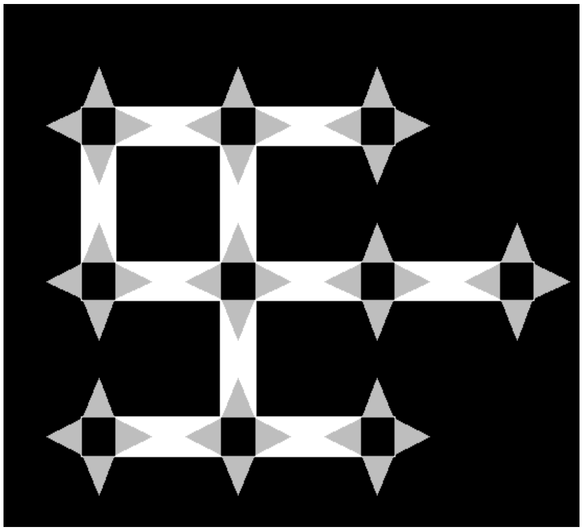


Figure 27: Map of the 10 corridor environment with 40 testable states and 5.32 bits of information available.

Eyelink II eye-tracker by SR Research was used to keep a record of the participant's eye movements throughout the experiment. For accuracy in data collection a chin rest was used to keep the participants head steady and at a uniform distance from the monitor screen. The chin rest was placed 100cm away from the monitor and the height was maintained at 33cm.

The *training* and *testing* session remained the same as in Experiment 1 and 2 (described in section 3.1.1 and 3.1.2). The difference was that the participant's head was placed on a chin rest and they wore the eye-tracker. The eyes were calibrated at the beginning of both the *training* and *testing* session. Successful completion of the *testing* phase ended the first session of the study. On the next day, for the second session the participant would start in the *testing* phase to make sure that they still remembered the environment. After passing the *testing* phase the participant started in the *experimental* phase (procedure explained in section 3.1.3).

7.1.1 Experimental Phase

We used the eye tracker during the *experimental* phase also. In this study the *experimental phase* was divided into two different versions that each participant completed. The '*Full*' version was like the *experimental* phase in Experiment 1, 2 and 3 (procedure explained in section 3.1.3). Each state was tested five times and the tested view was visible only for 1.5 seconds after which the map of the environment appeared on the screen. All of the landmarks were present in the environment.

In the '*Half*' version, the method remained the same but half of the landmarks were removed from the environment. Landmarks were removed by using the eye-tracker data from both the *training* and *testing* phase of session one. The total amount of time a participant spent looking at each landmark was calculated. From each tested view the total time spent looking at all visible landmarks was averaged. The tested views are the triangles on the map of the environment as seen in Figure 27[7]. The landmarks are placed along the corridors at a uniform distance and at the end of hallways as can be seen in Figure 28.

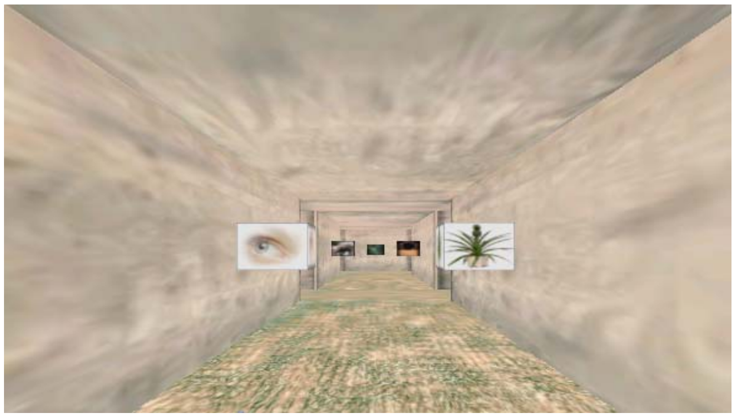


Figure 28: Sample view from an environment showing the landmarks visible at the end of the hallway and in the middle of the hallway.

Landmarks removed were either viewed more then the average amount of time or the ones that were viewed less then the average amount of time depending on the condition being tested in the *experimental phase*. Participants in the Low-Landmark-Viewing Group were tested in an environment that had the landmarks that were viewed less than average. The High-Landmark-Viewing Group was tested in an environment that had landmarks that had a viewing time more than average. Both groups were also tested in an environment with all the landmarks ('*Full*'). The order of '*Full*' vs. '*Half'* was counterbalanced across subjects.

[7] In the experiment we did not test the dead end views, the views of a single picture at the end of a corridor.

7.2 Results

Accuracy was calculated and compared for each participant in the '*Full*' and '*Half*' condition. Figure 29 shows the accuracy results for participants in Low-Landmark-Viewing Group and High-Landmark-Viewing Group.

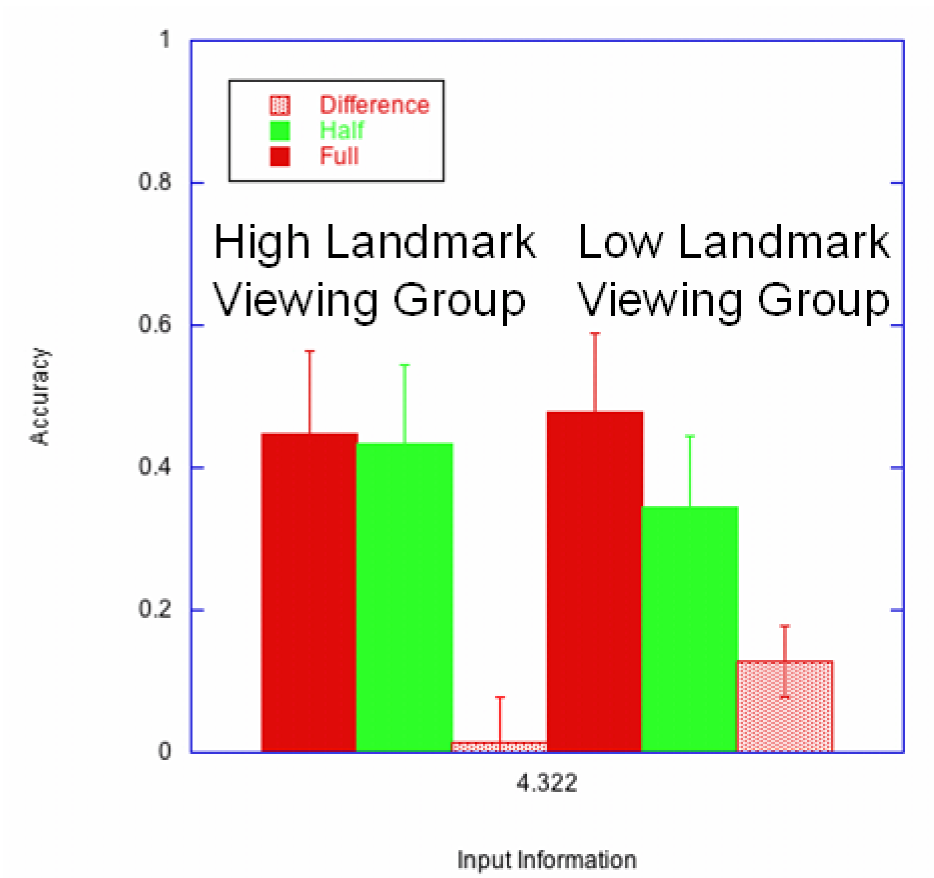


Figure 29: Plot shows the mean accuracy of participant's for each condition. The green marker shows the data for the participants in the 'Full' condition and the red marker shows the accuracy results from the 'Half' condition. High Landmark Viewing Group had the least viewed landmarks removed in the 'Half' condition. Low Landmark Viewing Group had the highly viewed landmarks removed in the 'Half' condition.

The accuracy results are the averaged accuracy of each participant according to the tested condition. There is a significant difference in performance for the Low-Landmark-Viewing Group when only the low-gaze-time landmarks were present in the '*Half*' condition (red marker in Figure 29) in comparison to when all of the landmarks were present ($t(6)=2.534$, $p=.04$) in the '*Full*' condition as seen by the green column in Figure 29 . Unlike the accuracy results for the Low-Landmark-Viewing Group, in the High-Landmark-Viewing Group we do not find a significant difference in accuracy when participants are viewing all of the visual landmarks in the '*Full*' condition as seen by the green column versus only the high-gaze-time landmarks ($t(6) = 0.229$, $p=.82$) in the '*Half*' condition as seen by the red column in Figure 29. This supports our earlier data in which about 1 to 1.5 bits of information were not being transferred to the *Cognitive Spatial Representation*. The removal of seldom viewed landmarks does not affect performance significantly. Memory is only storing pieces of the information available in an environment to aid in successful navigation.

Figure 30, shows the average amount of time all participants spent looking at landmarks. The landmarks are divided according to their position in the environment. The dark brown bars in Figure 30, shows the average amount of time that the participants spent looking at the landmark with the close up time being removed. Close up time would be when they were right in front of a particular landmark and unable to see any other landmarks. The light brown bars plots the average of the complete time (close up time included) spent looking at each landmark according to its position.

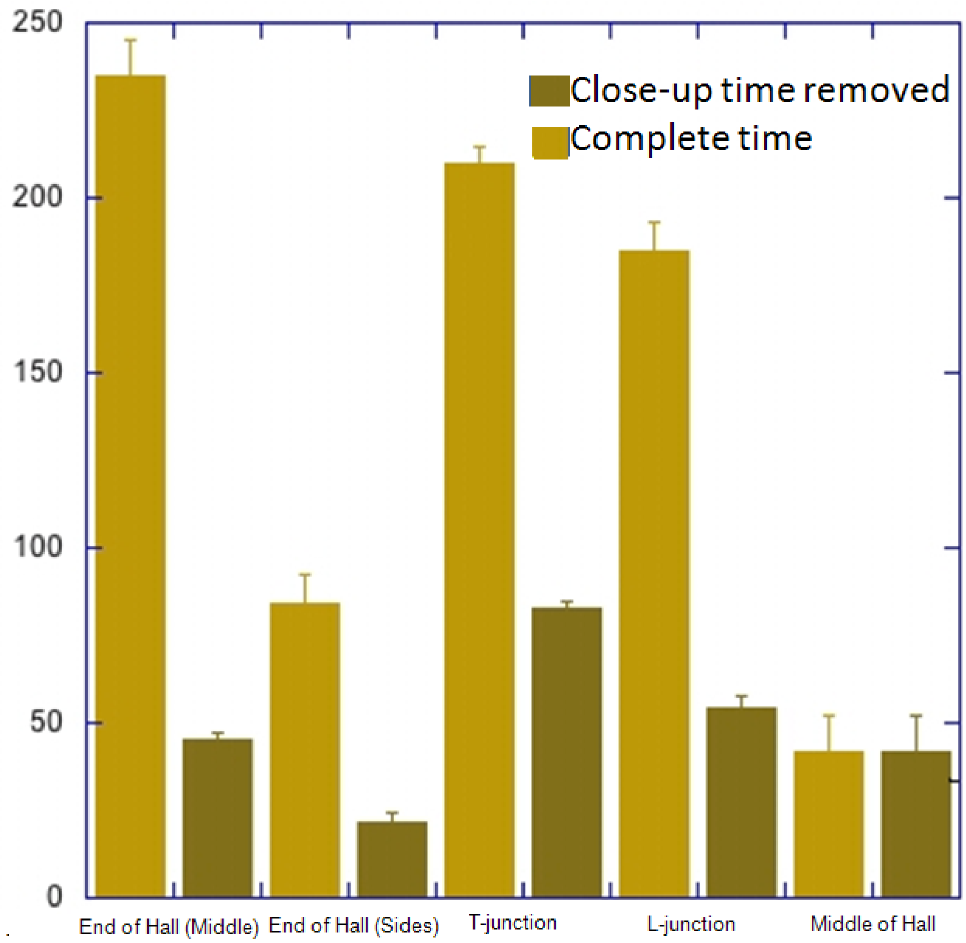


Figure 30: Plot shows the averaged time each participant spent looking at the landmarks. The landmarks are grouped together according to their location in the environment. 'End of Hall (middle)' are the landmarks in the center at the end of a corridor. 'End of Hall (sides)' are the landmarks located at the on the sides at the end of a corridor. 'T-junction' and 'L-junction' are the landmarks placed at the T-junctions and L-junctions in the environment. 'Middle' are the landmarks placed on either side in the middle of the corridor.

Figure 30 shows that the participants tend to spend more time looking at the landmarks at the end of corridors, T and L-junctions rather than landmarks located on the sides of corridors. Landmarks that are visible from several views can help an individual localize oneself in a number of views. Landmarks at the end of corridors would be visible from the most number of views. On the other hand learning a landmark that is visible from several views can also be confusing as an individual can make distance or orientation errors (left or right).

Figure 31a shows the landmarks (highlighted in red) that were removed most often for all seven or six of the participants in the 'half' experimental condition of the High-Landmark-Viewing Group. The landmarks in red are the ones that the majority of the participants spent the least amount of time looking at during the *testing* and *training phase*. Figure 31b shows the location of the landmarks that had the longest gaze time for most of the participants (all seven or six participants) during the *testing* and *training phase* and were hence removed for all participants during the 'half' *experimental phase* of the Low-Landmark-Viewing group.

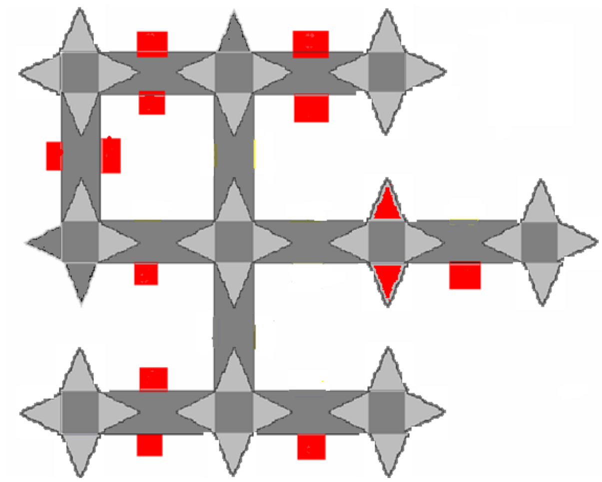


Figure 31a: Highlighted in red are the positions of the landmarks removed for almost all the participants in the High-Landmark-Viewing Group 'Half' Experimental Phase.

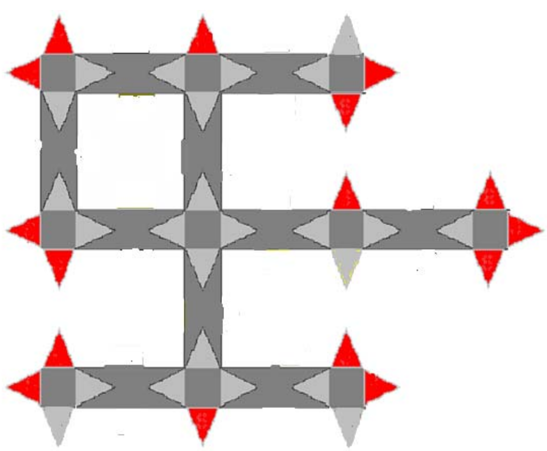


Figure 31b: Highlighted in red are the positions of the landmarks removed for almost all the participants in the Low-Landmark-Viewing Group 'Half' Experimental Phase.

Figure 30 and 31a &b show that there is a common pattern in how the participants distribute their gaze in the environment. The participants tend to spend more time looking at specific landmarks at decision points in the environment. The gaze pattern is not random without any common strategy to be found across participants. Landmarks at the end of corridors, L and T-junctions are the ones that get the majority of the gaze time. On the other hand there are other landmarks that are barely looked at when exploring and learning a new environment. Figure 32 shows gaze time on landmarks for two individual subjects from the experiment as an example of individual gaze distribution. This pattern of gaze was seen across all subjects: some landmarks are looked at a lot whereas others have very little to no gaze time.

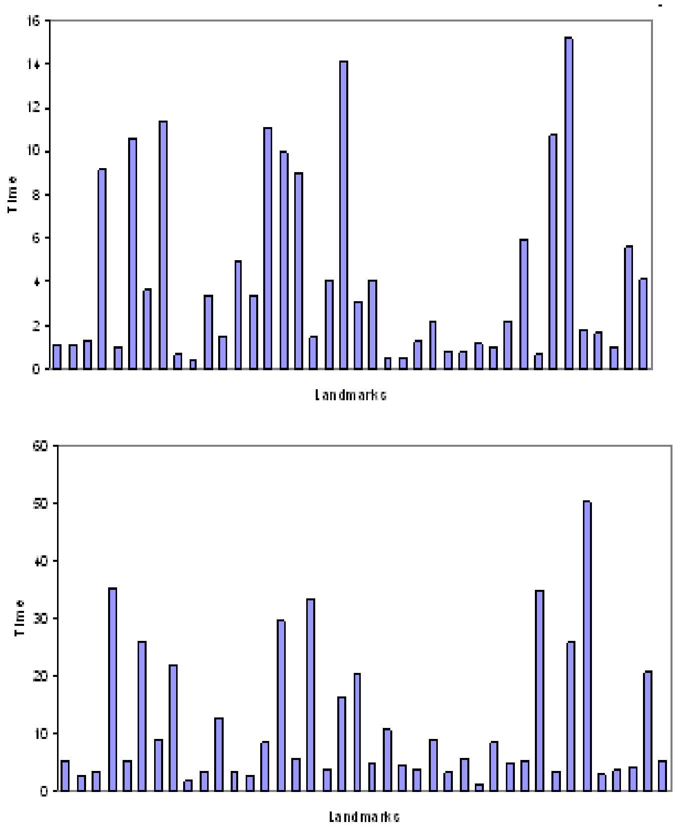


Figure 32: These plots show the amount of time (in seconds) two subjects from the study spent looking at each one of the 40 landmarks in the environment. X-axis shows each one of the landmarks and the y axis plots time in seconds.

The location of the object landmarks that would be removed for participants depending on whether they were tested in the Low-Landmark-Viewing Group or High-Landmark-Viewing Group is predictable. The High-Landmark-Viewing Group participants are using landmarks that can help an individual localize in multiple views from the environment. For example landmarks at the end of the hallway will be removed in the High-Landmark-Viewing Group as learning that one landmark can help an individual localize themselves all along the hall. The object landmark at the end of the hall can tell a participant which specific hall they are in, which direction they are facing in the hall (orientation) and the specific position in the hall by calculating the distance from the landmark. With limited memory space learning the end of the hall landmark can help an individual localize in several different views of the environment.

CHAPTER 8. EXPERIMENT FOUR (B)

As a follow up study for Experiment Four (A) we replicated the experiment using a smaller (5 corridor with 24 landmarks) and a larger (15 corridor with 56 landmarks) environment (same as the ones used in Experiment One and Three). Figure 33 show the map for the 5 and 15 corridor environment used.

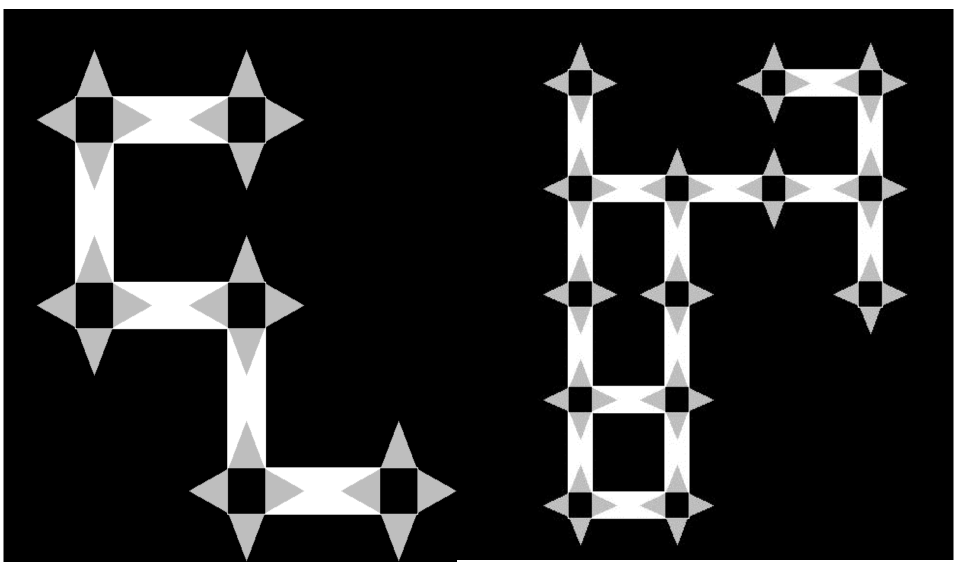


Figure 33: Map of 5 and 15 corridor environments used in Experiment 4B.

A smaller and a larger environment were tested to see how layout size affects gaze distribution. Are the landmarks being used in the same way and are the positions of the landmarks with the highest gaze timings remaining the same as found in Experiment Four (A)?

8.1 Procedure

The material used and procedure for Experiment Four (B) remained exactly the same as in Experiment Four (A) (described in section 7.1). Fourteen students from the University of Texas at Austin were paid $10 for their participation in the study. Participants were randomly assigned to be tested in the 5 corridor or 15 corridor environments. 4 participants each were randomly assigned to the High-Landmark-Viewing Group and the Low-Landmark-Viewing Group for the 15 corridor environment. 3 participants were randomly assigned to the High-Landmark-Viewing Group and Low-Landmark-Viewing Group for the 5 corridor environment.

8.2 Results

Figure 34 shows the accuracy results for the 5 corridor environment and Figure 35 shows the accuracy results for all participants in the 15 corridor environment. The red columns in both plots show the mean accuracy in the ‘Full’ condition when the participants were tested with all the landmarks visible in the view. The green columns show the mean accuracy for the ‘Half’ condition in which half of the landmarks were removed from the tested view. The dotted column shows the mean difference between the performance in the ‘Full’ and ‘Half condition.

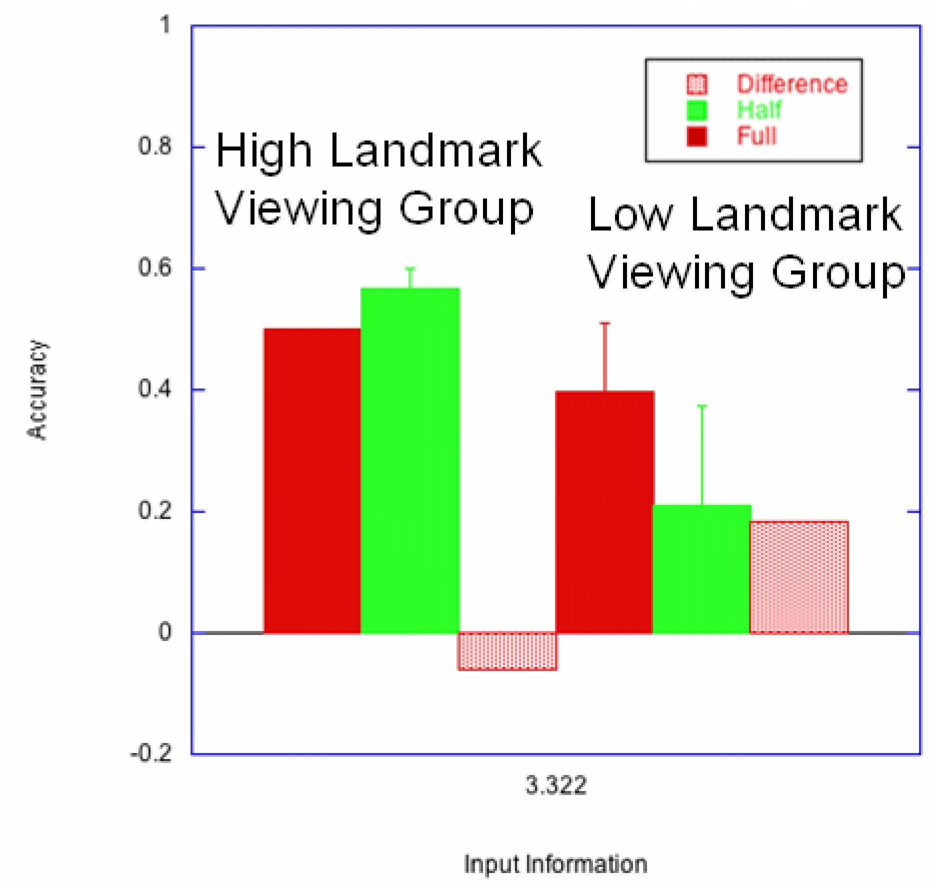


Figure 34: Plot shows the mean accuracy of participant’s who completed the 5 corridor environment. The green marker shows the data for the participants in the ‘Full’ condition and the red marker shows the accuracy results from the ‘Half’ condition. High Landmark Viewing Group had the least viewed landmarks removed in the ‘Half’ condition. Low Landmark Viewing Group had the highly viewed landmarks removed in the ‘Half’ condition.

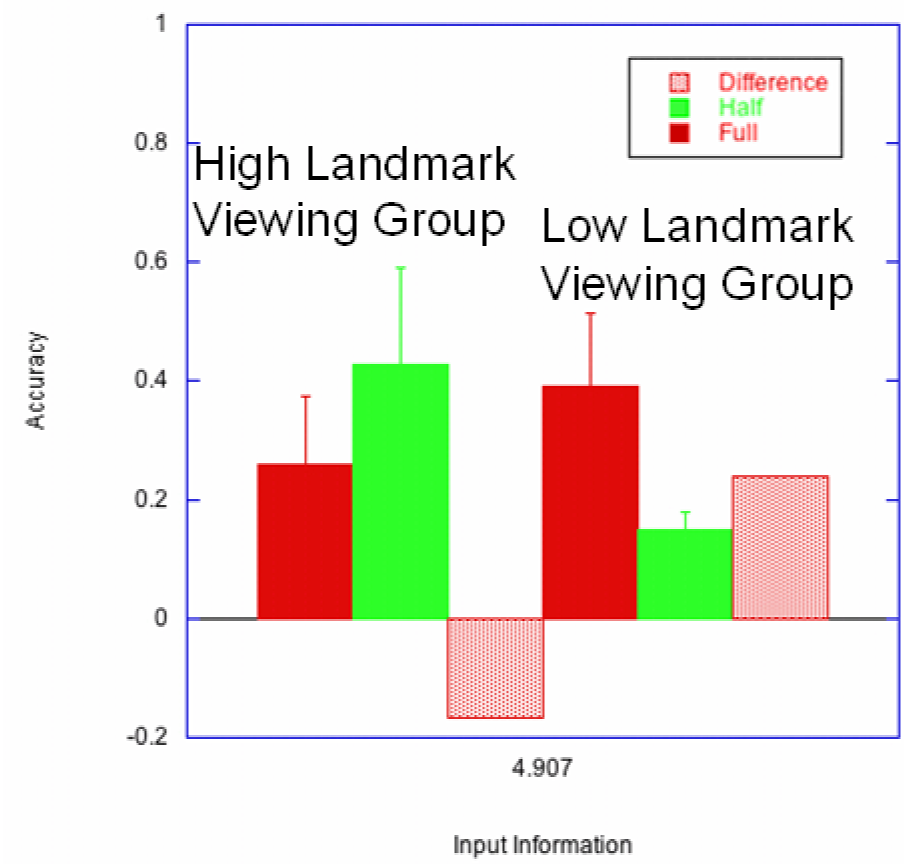


Figure 35: Plot shows the mean accuracy of participant's who completed the 15 corridor environment. The green marker shows the data for the participants in the 'Full' condition and the red marker shows the accuracy results from the 'Half' condition. High Landmark Viewing Group had the least viewed landmarks removed in the 'Half' condition. Low Landmark Viewing Group had the highly viewed landmarks removed in the 'Half' condition.

Figure 34 and 35 show that the accuracy results are consistent with Experiment 4(A). Participants clearly perform better in the 'Full' condition (5 corridor environment: t (2) = 3.541, p=.071) (15 corridor environment: t (3) = 1.837, p=.14) versus the 'Half' condition for the Low-Landmark-Viewing Group regardless of the size of the environment. When the seldom viewed landmarks are removed in the High-Landmark-Viewing Group performance actually improves in both environments for the 'Half' condition unlike in Experiment 4(A) where the difference between the 'Full' and 'Half' condition was minimal (5 corridor environment: t (2) = -1.99, p=.18) (15 corridor environment: t (3) = 3.341, p=.04).

Figure 36a shows the landmarks that were viewed for the least amount of time during the *training* and *testing phase* by either all three of the participants or two of the participants who were tested in the 5-corridor environment's High-Landmark-Viewing Group. Figure 36b shows the landmarks that were viewed the most by either all four of the participants or three of the participants and were hence removed during the 'Half' *experimental phase* for the Low-landmark-Viewing Group for the 5-corridor environment.

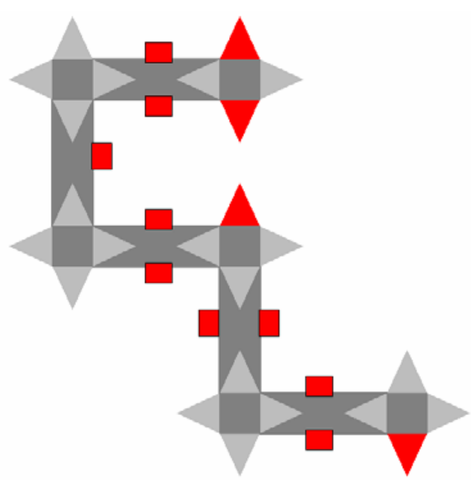


Figure 36a: Highlighted in red are the positions the landmarks that were removed for almost the participants in the High-Landmark Group, 'Half' experimental phase.

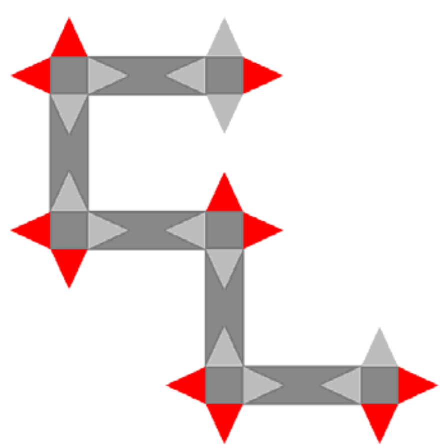


Figure 36b: Highlighted in red are the positions of of the landmarks that were removed for almost all of all the participants in the Low-Landmark-Viewing Viewing Group, 'Half' experimental phase.

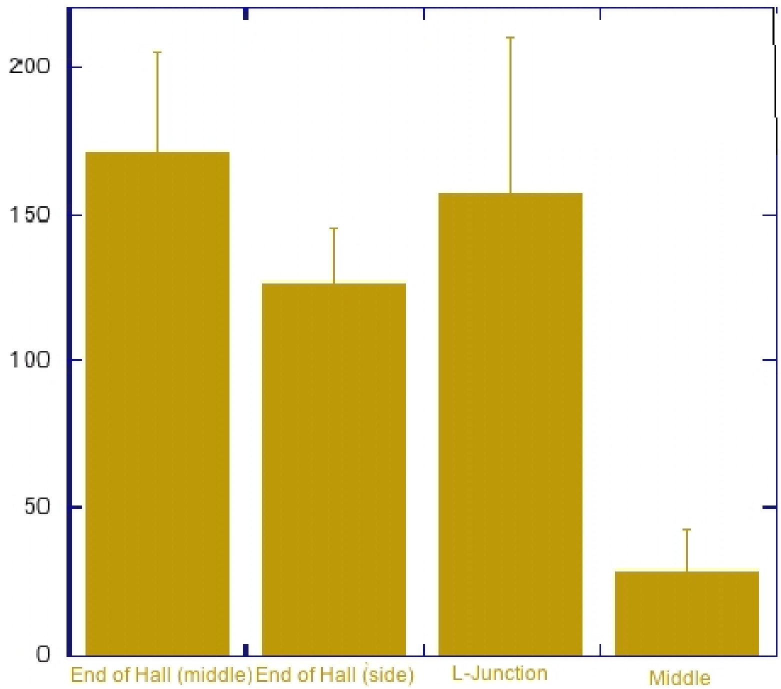


Figure 37: Plot shows the averaged time each participant spent looking at the landmarks in Environment 5. The landmarks are grouped together according to their location in the environment. 'End of Hall (middle)' are the landmarks in the center at the end of a corridor. 'End of Hall (sides)' are the landmarks located at the on the sides at the end of a corridor. 'L-junction' are the landmarks placed at the L-junctions in the environment. 'Middle' are the landmarks placed on either side in the middle of the corridor. There were no 'T-junctions' in this environment.

Figure 37 shows the average amount of time (seconds) that the participants spent looking at the landmarks. The landmarks are grouped according to their position in the environment. Environment 5 did not have any T-junctions. As seen for the 10 corridor environment participants spent more time looking at landmarks at the end of the hall (middle) and L-junctions. These are the landmarks that can be seen in the most number of views. Figure 38 shows sample gaze distribution from two participants in the study. Again there is a clear pattern visible in the gaze distribution of participants.

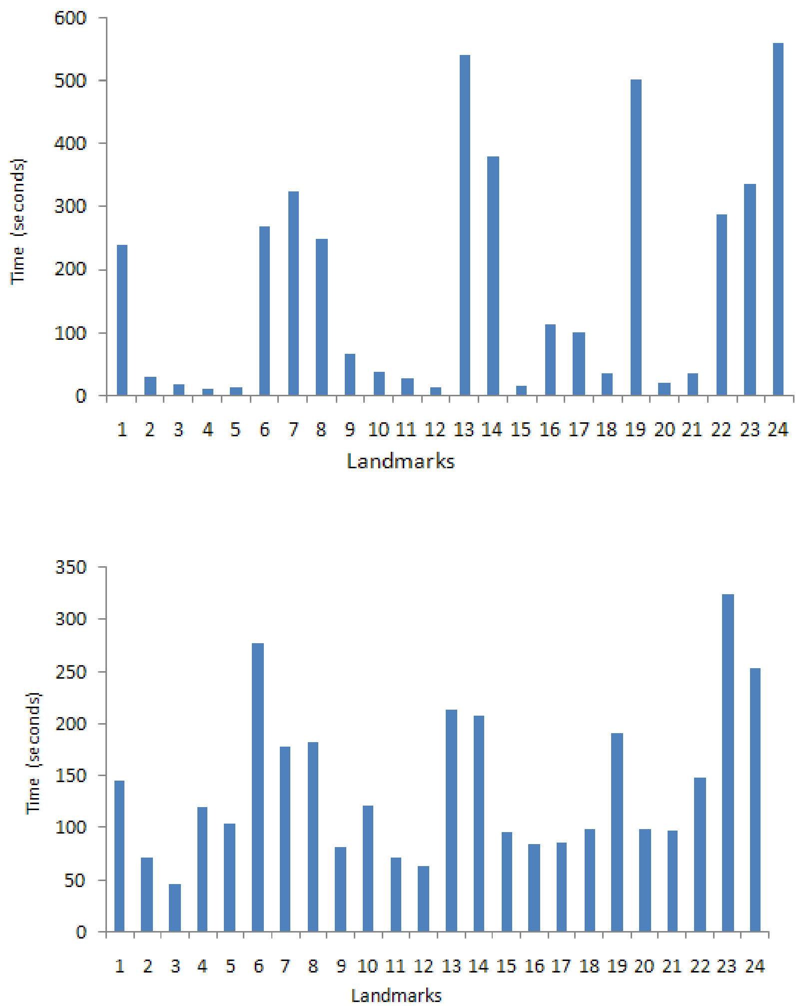


Figure 38: These plots show the amount of time (in seconds) two subjects from the study spent looking at each one of the 24 landmarks in the environment. X-axis shows each one of the landmarks and the y axis plots time in seconds.

Figure 39a shows the least viewed landmarks in the 15 corridor environment. Highlighted in red are the landmarks that were removed for all 5 or 4 of the participants in the High-Landmark-Viewing Group. Figure 39b shows the highly viewed landmarks for the Low-Landmark-Viewing Group that were removed for all 5 or 4 of the participants during the 'Half' *experimental phase*.

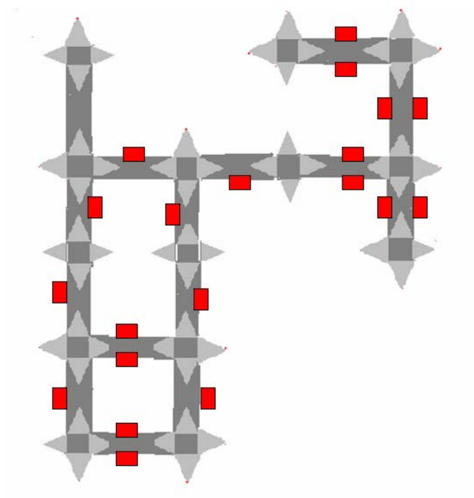


Figure 39a: Highlighted in red are the positions the landmarks that were removed for almost the participants in the High-Landmark Group, 'Half' experimental phase.

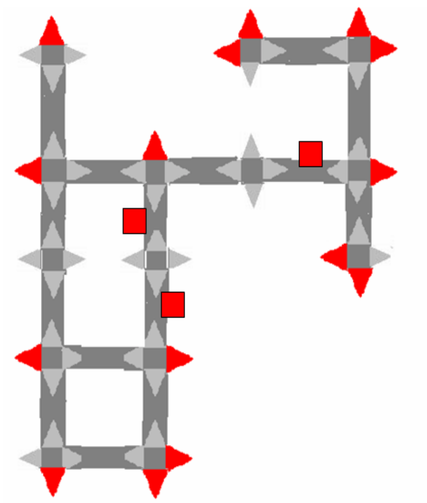


Figure 39b: Highlighted in red are the positions of of the landmarks that were removed for almost all of all the participants in the Low-Landmark-Viewing Viewing Group, 'Half' experimental phase.

Figures 36 and 39(a&b) are consistent with the Figure 31(a&b) from Experiment 4(A). The landmarks with the highest gaze times are located at the end of corridors, t-junctions or l-junctions. Landmarks placed in the middle of corridors tend to have minimum gaze times and are not encoded into the *Cognitive Spatial Representation.* Gaze distribution strategy remains the same across the different sized (5, 10 and 15 corridors) environments as seen in Figure 37 and 40.

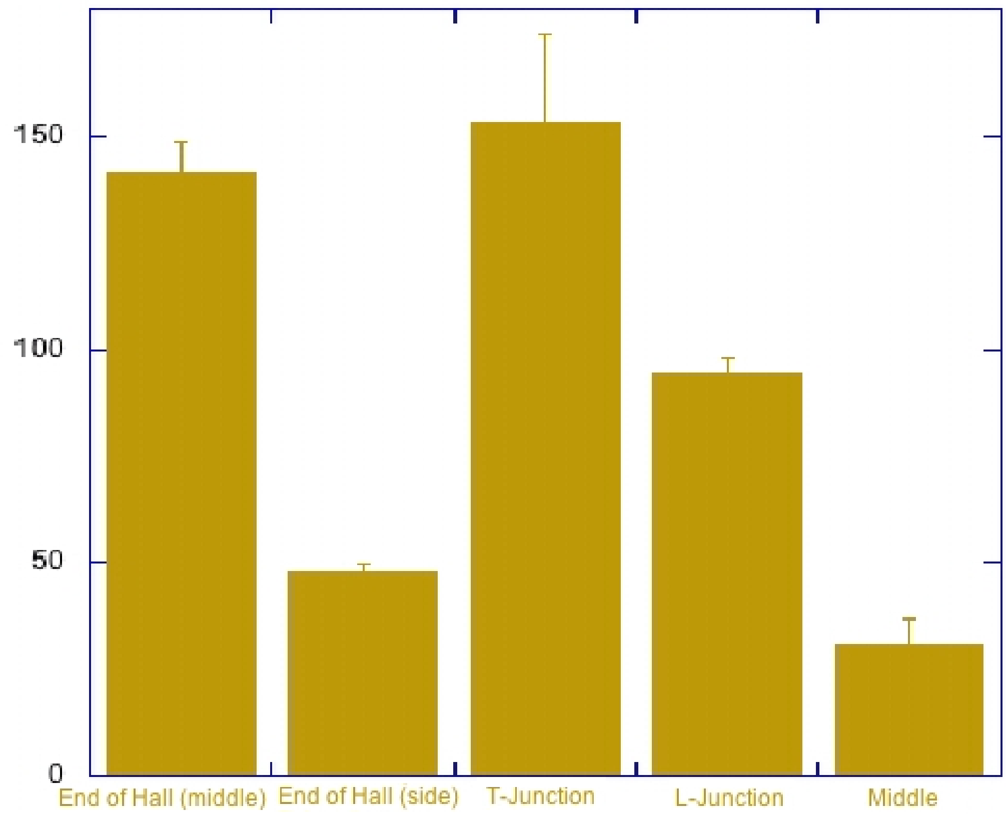


Figure 40: Plot shows the averaged time each participant spent looking at the landmarks. The landmarks are grouped together according to their location in the environment. 'End of Hall (middle)' are the landmarks in the center at the end of a corridor. 'End of Hall (sides)' are the landmarks located at the on the sides at the end of a corridor. 'T-junction' and 'L-junction' are the landmarks placed at the T-junctions and L-junctions in the environment. 'Middle' are the landmarks placed on either side in the middle of the corridor.

Figure 40 shows that the amount of time a person spends looking at a particular landmark is dependent on the location of the landmark in the environment. Gaze is not distributed equally amongst all object landmarks available in an environment. Landmarks placed at the end of hall (middle) and at T-junctions are the positions where participants spent the most of their gaze time when learning an environment. This is consistent with the results seen in the 5 and

10 corridor environment. The gaze distribution pattern remains consistent despite the change in size of the environment.

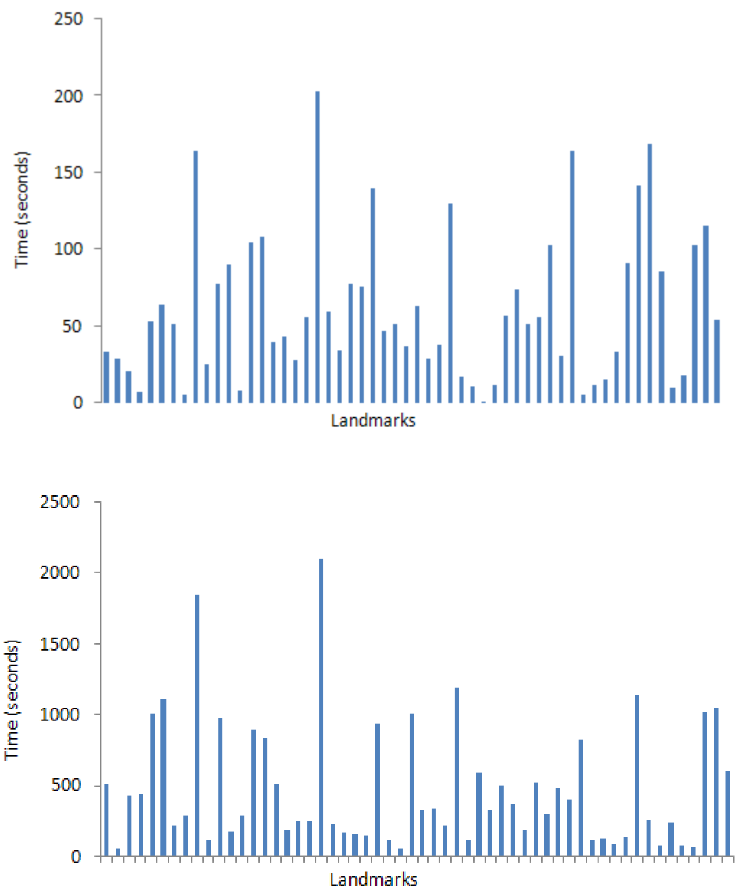


Figure 41: These plots show the amount of time (in seconds) two subjects from the study spent looking at each one of the 55 landmarks in the environment. X-axis shows each one of the landmarks and the y axis plots time in seconds.

Figure 41 shows examples of individual gaze distribution patterns of two participants from the study. As seen earlier the participants tend to spent spend a lot of gaze time at certain landmarks whereas on the other hand there are some object landmarks that are barely looked at when learning a new environment. Gaze distribution pattern remains consistent across all participants.

CHAPTER 9. GENERAL DISCUSSION

Stankiewicz et al. (2006) investigated human navigation skills when there is uncertainty about their current state in the environment and there is little visual information to help remove ambiguity. They used similar environments as used for our studies. They compared the efficiency of human navigation with that of the ideal-observer model that performed at optimal levels for each task. Results showed that participants became less efficient in their navigation abilities as the environments became larger which is consistent with our findings. The environments used in the Stankiewicz et al. (2006) studies were devoid of highly informative landmarks as in our work. In the fog condition their studies did show a decrease in efficiency when navigating but the difference between the fog and no fog condition was not significant.

Current research has looked at the amount of information being transferred to the human *Cognitive Spatial Representation*. The results suggest that there is no capacity limitation for up to 7.04 bits of information (132 states). There is a linearly increasing function for the bits of information being transferred in respect to the size of the environment. However, the transfer of information is not perfect. There is a loss of information regardless of the size of the environment. Interestingly, these results have been consistent over all the various testing conditions (trained participants, naïve participants, real world environment). Trained subjects also show an accuracy of about 50% regardless of the environment size. In all other conditions accuracy steadily decreases as the environments get bigger. In the real world environment the decrease in accuracy as the environment size increases is not as great as seen for the naïve subjects. The results support the view that people are not using all the information available to them. This would be a good strategy as information in a large-scale environment can be redundant. An efficient system would transfer the information that is highly informative without wasting resources on redundant information that adds little to the *Cognitive Spatial Representation*.

The limited view condition (fog) did not have a significant effect on the number of bits being transferred for participants who had learned the environment. Accuracy results for the fog condition with trained subjects started of well in the smaller sized environments but then steadily decreased as the environments got bigger. The mutual information results from the fog condition for naïve subjects show a significant decrease in the number of bits being transferred in comparison to the no-fog condition. Accuracy for this condition is very low regardless of the environment size. In a recent study by Foo et al. (2005) subjects performed navigation tasks in both a virtual "forest" environment where landmark information was plentiful, and a "desert" world where landmark information was sparse. Subjects could navigate in both environments, but could only choose novel routes or "shortcuts" in the "forest" environment. These results support the idea that people may actually need topological information for the development of their *Cognitive Spatial Representation*. In which case, people might not have developed a *Cognitive Spatial Representation* in environments, like open deserts or areas with fog, where only metrical information was available. The decline in accuracy of subjects in Experiment Two supports this idea. The naive subjects due to the lack of training do not seem to be able to form an accurate *Cognitive Spatial Representation* to aid them in navigation.

Follow up studies looked at what specific information is encoded within large-scale spaces (buildings or cities) that play an important role in way-finding and localization abilities. Results found that the impact of removing landmarks from an environment depended upon how the landmarks were used while the environment was being learned. Analyses have shown that people are not using all the information available in the environment as memory capacity is limited and not all the information available in an environment can be stored. Specific strategies are used in choosing the information to be stored to overcome memory limitations. The gaze distribution of participants is affected by the location of the landmark. Participants tend to spend more time looking at landmarks that can be seen from the most number of views when learning an environment. The *Spatial Cognitive Representation* of the participants encodes about half of the object landmarks in an environment. Removing the rest of the information with low gaze times does not affect performance as it is not encoded in the formation of the *Spatial Cognitive Representation*. Gaze distribution patterns remain consistent across different sized environments. The strategy for choosing the information to be encoded remains the same across different sized large scale environments.

Memory limitation is not only a problem for humans but even robotics's has to deal with the limited memory capacity of machines. A strategy for learning an environment without running out of memory capacity can help navigation skills in AI too. Navigation aids for individuals with low vision can also be made more usable for larger spaces if we know what specific information is the most beneficial to store. Results show that information at the center, at the end of corridors is the most likely to encoded while learning a new environment. Information in the middle of corridors has lower chances of being encoded as people spend less time looking at them.

References

Abu-Ghazzeh, T. M. (1996). Movement and way finding in the King Saud University built environment: a look at freshman orientation and environmental information. *Journal of Environmental Psychology, 16,* 303-318.

Aginsky, V., Harris, C., Resnick, R., & Beusmans, J. (1997). Two strategies for learning a route in a driving simulator. *Journal of Environmental Psychology, 17,* 317-331.

Cornell, E. H., Heth, C. D., & Alberts, D. M. (1994). Place recognition and way finding by children and adults. *Memory & Cognition, 22,* 633-643.

Cornell, E. H., Heth, C. D., & Rowat, W. L. (1992). Wayfinding by children and adults: Response to instructions to use look-back and retrace strategies. *Developmental Psychology, 28,* 328-336.

Cover, T., M., & Thomas, J., A. (2006). *Elements of Information Theory (2nd Ed.).*Wiley-Interscience Publication, New York.

Darken, R. P., & Sibert, J. L. (1996). Navigating in large virtual worlds. *International Journal of Human-Computer Interaction, 8(1),* 49-72.

Foo, P., Warren, W. H., Duchona, A., & Tarr, M. J. (2005). Do humans integrate routes into a cognitive map? map- versus landmark-based navigation of novel shortcuts. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31(2),* 195-215.

Franz, M., Sch¨olkopf, B., Mallot, H., & B¨ulthoff, H. (1998). Where did I take that snapshot. *Biological Cyberneticx, 79*, 191 – 202.

Gallistel, C. R. (1990) *The organization of learning.* Cambridge, MA: Bradford Books/MIT Press.

Gillner, S., & Mallot, H. A., (1998). Navigation and acquisition of spatial knowledge in a virtual maze. *Journal of Cognitive Neuroscience, 10(4),* 445-463.

Golledge, R. J. (1999). *Wayfinding behavior: cognitive mapping and other spatial processes.* Johns Hopkins University Press, Baltimore.

Judd, S. P. D., Dale, K., & Collett T.S. *On the fine structure of view-based navigation in insects. In Golledge (Eds.). Wayfinding behavior: cognitive mapping and other spatial processes.* Johns Hopkins University Press, Baltimore.

Hull, C. L. (1943). *Principles of behavior*. New York, NY: Appleton-Century-Crofts.

Kuipers, B. J., Byun, Y. T. (1991). A robot exploration and mapping strategy based on a semantic hierarchy of spatial representations. *Journal of Robotics and Autonomous Systems, 8*, 47-63

Kuipers, B. J. (1978). Modeling spatial knowledge. *Cognitive Science, 2*, 129-153.

Kuipers, B. J.. (2000). The spatial semantic hierarchy. *Artificial Intelligence, 119*, 191-233.

Kuipers, B. J.. (2001). The skeleton in the cognitive map: a computational hypothesis. *In J. Peponis, J. Wineman and S. Bafna (Eds.), Space Syntax: Proceedings of the Third*

International Symposium, Ann Arbor: A. Alfred Taubman College of Architecture and Urban Planning, University of Michigan, 10.1-10.7.

Mackay, D. J. C. (2003). *Information theory, inference and learning algorithms*. Cambridge University Press.

Mallot, H. A., Franz, M. O., Scholkopf, B., & Bulthoff, H. H. (1997). The view-graph approach to Navigation and Spatial Memory. In *ICANN* , 751-756. New York.

Mallot, H. A., & Gillner, S. (2000). Route navigating without place recognition: What is recognised in recognition-triggered responses? *Perception, 29,* 43-55.

May, A. J., Ross, T., & Byer, S. H. (2005). Incorporating landmarks in driver navigation system design: an overview of results from the REGIONAL project. *The Journal of Navigation, 58,* 47-65.

Miller, G. (1956). The magical number seven, plus or minus two: somem limits on our capacity for processing information. *The Psychological Review, 63 (2),* 81-97

Moar, I., & Carleton, L. R., (1982). Memory for routes. *Quarterely Journal of Experimental Psychology: Human Experimental Psychology, 34(A)*, 381-394.

Redish, A. D. (1999) *Beyond the Cognitive Map: From place cells to episodic memory*. Cambridge, MA, MIT Press

Reza, F. M. (1994). *An introduction to information theory.* New York, USA, Dover Publictaions.

Rodrigo, T., Chamizo, V. D., McLaren, I. P. L., & Mackintosh, N. J., (1997). Blocking in the spatial domain. *Journal of Experimental Psychology: Animal Behavior Processes, 23,* 110-118.

Ruddle, R. A., Payne, S. A., & Jones, D. M. (1997). Navigating buildings in "desk-top" virtual environments: Experimental investigations using extended navigational experience. *Journal of Experimental Psychology: Applied, 3,* 143-159.

Sanchez-Moreno, J., Rodrigo, T., Chamizo, V. D., & Mackintosh, N. J., (1999). Overshadowing in the spatial domain. *Animal Learning and Behavior, 27*, 391-398.

Schoelkopf, B., & Mallot, H. (1995). View-based cognitive mapping and path planning. *Adaptive Behavior*, *3,* 311-348.

Shannon, C. E. (1948). A mathematical theory of communication. *The Bell System Technical Journal, 27,* 379-423.

Shannon, C. E. (1993). The best detection of pulses. In *Collected Papers,* ed. By N. J. A. Sloane and A. D. Wyner, 148-150. New York: IEEE Press.

Shannon, C. E., & Weaver, W. (1949). *The mathematical theory of communication.* University of Illinois Press, Urbana.

Stankiewicz, B. J., & Kalia, A. (2007). Acquisition of structural versus object landmark knowledge. *Journal of Experimental Psychology: Human Perception and Performance, 33(2),* 378-390.

Stankiewicz, B. J., Legge, G. E., Mansfield, J. S., & Schlicht, E. J. (2006). Lost in virtual space: Studies in human and ideal spatial navigation. *Journal of Experimental Psychology: Human Perception and Performance, 32(3),* 688-704.

Stevens, A., & Coupe, P. (1978). Distortions in judged spatial relations. *Cognitive Psychology, 13*, 422-437.

Siegel, A. W., & White, S. H. (1975). The development of spatial representations of large-scale environments. *In H. W. Reese (Eds.), Advances in Child Development and Behavior, 10*. New York: Academic Press.

Tolman, E. C. (1948). Cognitive maps in rats and men. *The Psychological Review, 55(4)*, 189--208.

Tolman, E. C., Ritchie, B. F., & Kalish, D. (1946). Place versus response learning. *Journal of Experimental Psychology, 36*, 224-229.

Tversky, B. (1993). Cognitive maps, cognitive collages, and spatial mental models. In A. Frank & Campari (Eds.), Spatial information theory: A theoretical basis for gis, cosit '93 (Vol. 716, p. 14-24). Berlin: Springer-Verlag.

Vinson. G., N. (1999). Design guidelines for landmarks to support navigation in virtual environments. In Proceedings of the *Annual ACM SIGCHI 1999 Conference*, 15-20.

VITA

Sahar Nadeem was born in Karachi, Pakistan, on December 22, 1981, the daughter of Ansa and Nadeem Ahmed Afghan. After completing her secondary education at Karachi Grammar School, Karachi, Pakistan, in 2000, she graduated from Roanoke College in May 2004. In August, 2004, she entered the Graduate School at The University of Texas at Austin.

Permanent Address: 75/2/1, 15th Street,

Phase-6, D.H.A.

Karachi-75500, Pakistan

This dissertation was typed by the author